DOCUMENT RESUME

ED 097 025 IR 001 212

TITLE	Abstracts of Research, July 1973 through June 1974.
INSTITUTION	Ohio State Univ., Columbus. Computer and Information Science Research Center.
REPORT NO	OSU-CISRC-6/74
PUB DATE	74
NOTE	94p.
EDRS PRICE	MF-$0.75 HC-$4.20 PLUS POSTAGE
DESCRIPTORS	*Abstracts; Annual Reports; *Artificial Intelligence; Computer Programs; Computers; *Computer Science; Cybernetics; Information Networks; *Information Processing; Information Retrieval; *Information Science; Information Storage; Information Systems; Linguistics; Mathematical Applications
IDENTIFIERS	*Ohio State University

ABSTRACT

Abstracts of research papers in the fields of computer and information science are given; 72 papers are abstracted in the areas of information storage and retrieval, information processing, linguistic analysis, artificial intelligence, mathematical techniques, systems programing, and computer networks. In addition, the Ohio State University Computer and Information Science Research Center is described. The abstracts are indexed by subject and investigator. (WH)

ED 09702

CISRC

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ABSTRACTS OF RESEARCH

July 1973 through June 1974

COMPUTER and INFORMATION SCIENCE RESEARCH CENTER

The Ohio State University
Columbus, Ohio 43210

IR 00 1 814

FOREWORD

The Computer and Information Science Research Center at The Ohio State University is a research organization which consists primarily of faculty, staff and graduate students of the Department of Computer and Information Science. Some of the research activities are performed in conjunction with other University departments as well as off-campus organizations.

This publication contains the abstracts of research which has been carried on during the 1973-74 academic year. This research has been supported in part by grants from governmental agencies as well as by The Ohio State University. Sponsorship with government agencies and with other units on the campus is identified at the end of an abstract.

A bibliography of the research reports published by the Center is included in this publication as Appendix F. Copies of some of these reports are still available on a complimentary basis from the Computer and Information Science Research Center, The Ohio State University, 2024 Neil Avenue, Columbus, Ohio, 43210. Titles with PB numbers may be obtained from The National Technical Information Center, The U.S. Department of Commerce, 5285 Port Royal Road, Springfield, Virginia, 22151, in paper copy, magnetic tape, or microfiche. There is a nominal charge for their service.

Marshall C. Yovits
Director, Computer and
Information Science Research Center

TABLE OF CONTENTS

FOREWARD ii

I. THE OHIO STATE UNIVERSITY COMPUTER AND INFORMATION SCIENCE RESEARCH CENTER

ORGANIZATIONAL STRUCTURE 1

OBJECTIVES OF THE CENTER 1

SCOPE OF THE PROGRAM 2

FACILITIES 2

ACADEMIC PROGRAMS IN COMPUTER AND INFORMATION SCIENCE 4

Organization of The Department of Computer and Information Science 4
Objectives of The Department 4
Growth of The Department of Computer and Information Science 5
Undergraduate Programs 6
Graduate Programs 6
Course Offerings 7
Faculty 8

INTERACTION WITHIN THE UNIVERSITY 8

INTERACTION WITHIN THE INFORMATION SCIENCE COMMUNITY 8

DOCTOR OF PHILOSOPHY DEGREE 9

RECENT TECHNICAL REPORTS 9

II INFORMATION STORAGE AND RETRIEVAL

ALGORITHMIC CONTROL OF INDEX TERM SPECIFICITY IN DOUBLE-KWIC COORDINATE INDEXING SYSTEMS A. E. Petrarca, W. M. Lay, W. S. Stalcup 11

AUTOMATIC INDEXING OF OUTPUTS FROM RETROSPECTIVE SEARCHING OF COMPUTER-READABLE FILES A. E. Petrarca, J. H. Hsu 11

AUTOMATIC INFORMATION SYSTEMS ANALYSIS THROUGH EVALUATION OF SIMULATED PERFORMANCE T. DeLutis, J. S. Chandler 12

AUTOMATIC VOCABULARY CONTROL IN PRINTED INDEXES GENERATED FROM NATURAL-LANGUAGE TEST A. E. Petrarca, W. S. Stalcup 12

DECISION THEORETIC APPROACH TO AUTOMATIC DOCUMENT CLASSIFICATION L. J. White, G. Kar 13

INFORMATION PROCESSING SYSTEMS SIMULATION T. DeLutis 13

N-GRAM KEYTERMS IN AUTOMATIC DOCUMENT CLASSIFICATION S. E. James 14

PERFORMANCE EVALUATION OF DATA BASE MANAGEMENT SYSTEMS T. DeLutis, J. Aitken 14

A QUANTITATIVE MEASURE OF CONCEPT SCATTERING FOR EVALUATION OF AUTOMATIC VOCABULARY CONTROL TECHNIQUES A. E. Petrarca, W. S. Stalcup 15

A SEQUENTIAL APPROACH TO AUTOMATIC DOCUMENT CLASSIFICATION L. J. White, S. E. James, G. P. Jones, G. Kar, D. E. Westbrook 16

III. HUMAN INFORMATION PROCESSING

HUMAN EXTRAPOLATION OF STRINGS GENERATED BY ORDERED CYCLIC FINITE STATE GRAMMARS J. L. Beug, R. Ernst 17

USE OF PROBABILISTIC AUTOMATA AS MODELS OF HUMAN PERFORMANCE S. P. Hepler, R. Ernst 17

IV. INFORMATION ANALYSIS

A GENERAL THEORY OF INFORMATION FLOW AND ANALYSIS M. C. Yovits, J. G. Abilock 19

V. LINGUISTIC ANALYSIS

PARALLEL RECOGNITION OF FORMAL LANGUAGES J. Rothstein, M. Moshell 21

ON THE SYNTACTIC STRUCTURES OF UNRESTRICTED GRAMMARS H. W. Buttelmann 21

GENERALIZED FINITE AUTOMATA ON UNRESTRICTED SYNTACTIC STRUCTURES H. W. Buttelmann 22

A FORMAL THEORY OF THE SYNTAX, SEMANTICS AND TRANSLATION OF PHRASE-STRUCTURE LANGUAGES H. W. Buttelmann 22

SYNTAX-DIRECTED AND SEMANTIC-DIRECTED TRANSLATION OF PHRASE STRUCTURE LANGUAGES H. W. Buttelmann 22

SEMANTIC-DIRECTED TRANSLATION--ANOTHER APPROACH H. W. Buttelmann, A. Pyster 23

STUDIES IN DECIDABILITY, EXISTENCE, AND EFFECTIVENESS OF TRANSLATIONS ON PHRASE-STRUCTURE LANGUAGES H. W. Buttelmann, R. Krishnaswamy 23

SOME PROPERTIES OF SYNTAX-DIRECTED TRANSLATIONS H. W. Buttelmann, F. J. Dickey 23

THE ROLE OF THE NONTERMINAL IN LANGUAGE GENERATION H. W. Buttelmann, A. Pyster 24

VI. ARTIFICIAL INTELLIGENCE

DECISION THEORETIC MODELS FOR INVARIANT PATTERN RECOGNITION L. J. White, S. N. Srihari 25

DECISION THEORY FOR PATTERN RECOGNITION B. Chandrasekaran, A. K. Jain 25

DISTANCE FUNCTIONS FOR INDEPENDENT MEASUREMENTS AND FINITE SAMPLE SIZE B. Chandrasekaran, A. K. Jain 26

FINITE MEMORY DECISION THEORY B. Chandrasekaran, C. C. Lam 26

GRAMMATICAL INFERENCE PROBLEM L. J. White, D. A. Marik 26

INTERACTIVE GENERATION OF HUMAN FACIAL IMAGE ON CRT USING A HEURISTIC STRATEGY B. Chandrasekaran, M. L. Gillenson 27

AN INTERACTIVE SYSTEM FOR AUTOMATIC PROGRAM SYNTHESIS A. W. Biermann, R. Krishnaswamy, F. Petry, B. Chandrasekaran 27

LOW-FREQUENCY RADAR AIRCRAFT DETECTION AND CLASSIFICATION L. J. White 28

NUMBER TREES, SEMIGROUPS, AND FORMAL LANGUAGES J. Rothstein 28

VII. INFORMATION PROCESSES IN PHYSICAL, BIOLOGICAL AND SOCIAL SYSTEMS

ON THE ROLE OF POST HOC ERGO PROPTER HOC IN SCIENTIFIC METHODOLOGY AND ARTIFICIAL INTELLIGENCE J. Rothstein 30

VIII. MATHEMATICAL TECHNIQUES

AN ALGORITHM FOR SUB-OPTIMAL SET COVERS WITH CONFIDENCE MEASURES OF OPTIMALITY L. J. White, M. E. Doherty 31

BANDWIDTH REDUCTION OF SPARSE MATRICES BY ROW AND COLUMN PERMUTATIONS D. S. Kerr, L. J. White, P. T. R. Wang 31

BINARY STRINGS AND TOPOLOGY J. Rothstein 31

COMPARISON OF ROUTINES FOR SOLVING NONLINEAR SYSTEMS OF EQUATIONS D. S. Kerr, L. J. White, D. L. Kalmey 32

MINIMUM-REDUNDANCY VARIABLE LENGTH CODING L. J. White, V. Santhanam 33

SIEVES AND CONTEXT SENSITIVE LANGUAGE J. Rothstein 33

SOME L^P-NORM MAXMIN PROBLEMS: ALGORITHMS AND APPLICATIONS L. J. White, B. Dasarathy 33

A STUDY OF THE OPTIMAL PLACEMENT OF KNOTS FOR CUBIC SPLINE APPROXIMATIONS L. J. White, S. Muftic, P. Wong 34

TOWARD AN ARITHMETIC FOR CELLULAR AUTOMATA AND PARALLEL COMPUTATION J. Rothstein 34

TURING UNIVERSALITY AND PARALLELISM OF GROUPOID STRING ALGORITHMS J. Rothstein 35

IX. SYSTEMS PROGRAMMING

CONTEXT PROTECTION AND CONSISTENT CONTROL IN DATA SECURE SYSTEMS D. K. Hsiao, C. J. Nee, D. Wong 36

A CROSS-ASSEMBLER FOR THE MICRODATA 1600/21 COMPUTER SYSTEM D. K. Hsiao, T. Rodeheffer 36

CRYPTOGRAPHIC TECHNIQUES AND DATA SECURITY IN DATA BASE SYSTEMS E. Gudes, H. S. Koch, F. Stahl 37

A DATA BASE MANAGEMENT SYSTEM WITH GENERALIZED DATA STRUCTURE, ACCESS AND UPDATE H. S. Koch, B. J. Kammerer, R. Knablein, H. Sieberth 37

DATA SECURE SYSTEM ARCHITECTURE D. K. Hsiao, R. Baum 38

A DATA SECURE SYSTEMS MODEL D. K. Hsiao, E. J. McCauley 39

DESIGN AND IMPLEMENTATION OF A DATA BASE PROTECTION MECHANISM D. K. Hsiao, N. Kaffen 39

DESIGN AND IMPLEMENTATION OF A METHOD FOR CONTROLLING COMPUTER PROGRAM-POSTMORTEM DIAGNOSTIC OUTPUT R. F. Mathis, M. S. Friedman 39

THE DEVELOPMENT OF GUIDELINES FOR DATA BASE REORGANIZATION H. S. Koch 40

FILE PARTITIONING AND RECORD PLACEMENT IN ATTRIBUTE-BASED FILE ORGANIZATIONS D. K. Hsiao, F. Manola, E. J. McCauley 41

GRAPHICAL MODEL OF FUNCTIONAL DEPENDENCIES IN A RELATIONAL DATA BASE D. S. Kerr 41

IMPROVED ERROR DIAGNOSTICS FOR IBM SYSTEM /370 OS/MVT ASSEMBLER PROGRAM DUMPS R. F. Mathis, B. M. Kirsch 41

INVESTIGATION OF CRYPTOGRAPHIC TECHNIQUES FOR COMPUTER SECURITY AND PRIVACY F. Stahl 41

INVESTIGATION OF PATTERN RECOGNITION TECHNIQUES AS APPLIED TO PROBLEMS OF COMPUTER SECURITY F. Stahl 42

MAPPING FUNCTIONS FOR INDEPENDENT DATA DEFINITION H. S. Koch 42

A MODEL FOR ACCESS AND ACCESS CONTROL IN DATA BASE MANAGEMENT SYSTEMS D. K. Hsiao, F. Manola 42

PAGINATION OF PROGRAMS FOR VIRTUAL MEMORY SYSTEMS C. R. Foulk, D. H. Y. Su 43

PARALLELISM IN COMPUTER PROGRAMS C. R. Foulk, O. C. Juelich 43

THE PL/X SYSTEMS PROGRAMMING LANGUAGE D. K. Hsiao, B. J. Kammerer, W. Young 44

PROVING PROGRAM CORRECTNESS ON COMPLEX FILE STRUCTURES D. K. Hsiao, B. Horger 44

A SMOOTHNESS INDEX FOR COMPUTER PROGRAM FLOWGRAPHS C. R. Foulk, O. Longe 45

TEACHING DEBUGGING R. F. Mathis, C. Hasbrouck 45

USE OF AN ERROR ANALYSIS CONTROL BLOCK (EACB) TO CONTROL PROMPTING IN A POSTMORTEM DEBUGGING SYSTEM R. F. Mathis, R. J. Atwell, Jr. 46

X. COMPUTER NETWORKS

DISTRIBUTED OPERATING SYSTEMS FOR COMPUTER NETWORKS M. T. Liu, S. Fournier 47

EMULATION OF COMPUTER NETWORKS BY MICROPROGRAMMABLE MICROCOMPUTERS M. T. Liu, D. Cohen 47

LOCAL COMPUTER NETWORK D. K. Hsiao, B. Farah, T. Rodeheffer 48

MIXED DATA/VOICE TRANSMISSION SYSTEM FOR COMPUTER COMMUNICATIONS M. T. Liu, J. T. Wang 48

OPTIMAL PROCESS ALLOCATION IN DISTRIBUTED HETEROGENEOUS COMPUTER NETWORKS M. T. Liu, T. T. Cheng 49

VARIABLE-LENGTH MESSAGE TRANSMISSION FOR DISTRIBUTED LOOP COMPUTER NETWORKS M. T. Liu, C. C. Reames 49

XI. JOINT PROGRAMS

INVESTIGATION OF COMPUTER SCIENCE EDUCATION IN LATIN AMERICA F. Stahl, A. Pumarada 50

APPENDIXES

A COMPUTER AND INFORMATION SCIENCE COURSE LISTING BY NUMBER AND TITLE 51

B COMPUTER AND INFORMATION SCIENCE FACULTY 55

C COMPUTER AND INFORMATION SCIENCE SEMINAR SERIES 59

D RELATED ACTIVITIES OF THE STAFF OF COMPUTER AND INFORMATION SCIENCE RESEARCH CENTER 62

E PUBLICATIONS OF THE COMPUTER AND INFORMATION SCIENCE RESEARCH CENTER STAFF 69

F TECHNICAL REPORT SERIES 73

INDEXES

INVESTIGATOR INDEX 78

SUBJECT INDEX 81

I. THE OHIO STATE UNIVERSITY COMPUTER AND INFORMATION SCIENCE RESEARCH CENTER

ORGANIZATIONAL STRUCTURE

The Computer and Information Science Research Center at The Ohio State University is an interdisciplinary activity involving the faculty, staff and graduate students from the Department of Computer and Information Science. Some of the research activities are performed in conjunction with other University departments as well as off-campus organizations. The Center also interacts closely with Battelle Memorial Institute, Chemical Abstracts Service, and the Ohio College Library Center, which are adjacent to the Ohio State Campus, as well as with a number of other organizations located in Columbus, Ohio which are engaged in computer and information science research activities, such as Bell Laboratories, Western Electric Corporation, and Rockwell International Corporation. Although the Research Center and the Department utilize many of the same personnel and have the same Director, they are two separate and distinct entities.

OBJECTIVES OF THE CENTER

The Computer and Information Science Research Center has the following objectives: (1) to develop a broad research program in computer and information science; (2) to develop, test, and evaluate practical applications of research in computer and information science; (3) to coordinate and integrate these functions with an academic program in computer and information science at The Ohio State University, as well as with other disciplines at the University. The Center is a focal point for a number of applied information processing activities on the campus.

SCOPE OF THE PROGRAM

The program in computer and information science at The Ohio State University has been defined broadly to encompass most of the analytical activities frequently considered to be part of this discipline. This approach has been chosen because it is felt that in order to generate the needed concepts, foundations, and generalized techniques, it is necessary to examine analytically a number of different areas of computer and information science. In this way a firm empirical and theoretical foundation may be established for generalized computer and information systems. Such a view commits the program not only to the theoretical study of computer and information systems, but also to the study of their realization and their impact on the user.

Those areas of study which are emphasized both in the academic program and in the research activities of the Center are as follows:

1. General theory of information
2. Information storage and retrieval
3. Theory of automata and theory of computation
4. Artificial intelligence
5. Pattern recognition
6. Computer programming, including systems programming
7. Theory and processing of programming languages
8. Digital computer architecture and organization
9. Numerical analysis
10. Man-machine interaction and systems
11. Formal and computational linguistics
12. Management information and systems
13. Biological information processing
14. Social, economic, and psychological aspects of information production, processing, and use.

FACILITIES

The Computer and Information Science Research Center has a Digital Equipment Corporation DEC System-10 computer which was obtained with the assistance of a grant from the Office of Computing Activities of the National Science Foundation. This computer is a moderate size flexible time-sharing computer which is dedicated to research and education in the field of computer and information science. This computer provides valuable hands-on experience for the faculty and students of the Computer and Information Science Research Center and permits research activities involving non-standard and innovative applications of computers of both a hardware and software nature. Some of these research activities which are currently underway are:

1) Experimental and developmental research in time-sharing and multiprogramming systems.
2) Complex systems simulation research using graphical display devices.

3) On-line information retrieval systems studies.
4) Hardware modification and interface studies.
5) Software modification and development (e.g., PL/1 and COBOL compilers).
6) Man-machine interaction and psychophysical experiments.
7) Pattern recognition studies.
8) Computer simulation of language learning.
9) Speech analysis and synthesis.
10) Analysis and synthesis of human locomotion.
11) Computer data base systems.

The Research Center also has access to the University Computer Centers. They are: Instruction and Research Computer Center; Hospital Computer Center; and University Systems Computer Center.

Included in these centers are an IBM 370/165, an IBM 370/158, an IBM 1620 for plotting, and an IBM 7 for analog to digital conversion, as well as several IBM 1130 machines, and a number of remote terminals.

Many specialized facilities and laboratories of the University are also available to the staff and students of the Research Center. Some of these are the Office of Computer Assisted Instruction, the Institute for Research in Vision, Telecommunications Center, Behavioral Science Laboratory, Listening Center, Communications and Control Systems Laboratory, Mershon Center for Education in National Security, and many others.

The Research Center also interacts with the Ohio College Library Center which is administratively independent of the University. The Center was formed by the Ohio College Association, and operates a common computerized library network connecting the Ohio colleges and universities (both private and state assisted) and many points outside the state of Ohio. These include eight regions with 100 colleges and universities, and 19 public and special libraries participating in the system. Two well established national information systems have units on the campus of The Ohio State University. The MEDLINE system is an automated on-line service to access medical journals of the previous three year period. The ERIC (Educational Research Information Center) system is an automated batch system to access research reports and journal literature in the field of education. These systems are available to the staff and students. Interaction has also been initiated with The Academy for Contemporary Problems, an organization formed jointly by The Ohio State University and Battelle Laboratories, Columbus.

The University has established a university-centered-information system. The information system, called the Mechanized Information Center (MIC) operates as a department of the University Libraries. MIC has developed a multi-disciplinary batch-mode information system from machine-readable data bases primarily for the campus scientific community. The interface to the MIC system is decentralized as much as possible through the existing system of twenty-three libraries around the campus which serve specialized publics. MIC acquires

data bases from commercial sources, as well as from professional societies and governmental agencies. Research activities in MIC are directed toward improving the services of MIC to its users through software refinement and development. This center works closely with the staff and students of the Department and the Center.

The main service now offered by MIC is a computer-based, current awareness system for helping faculty and students keep up with published information, such as;

1. articles in the current issues of more than 3,200 journals
2. unclassified government reports included in the most recent issue of Government Reports Announcements
3. books cataloged by the Library of Congress
4. current conference papers presented at professional meetings and conferences held throughout the world.

In one year's time, a person, through his profile, can scan approximately 550,000 items, including 400,000 articles, 50,000 government reports, and 100,000 books. In addition, retrospective service (1968 to the present) is available, which covers several million journal citations.

ACADEMIC PROGRAMS IN COMPUTER AND INFORMATION SCIENCE

The program at The Ohio State University emphasizes education, research, and the professional practice and application of computer and information science. The department offers undergraduate and graduate degrees through the Ph.D. Statistics showing the growth of the department are found in Table 1.

Organization of The Department of Computer and Information Science

The Department of Computer and Information Science is a separate academic unit located administratively in the College of Engineering, operating in part as an interdisciplinary program with the cooperation of many other departments and colleges throughout the University.

Objectives of The Department

The program at The Ohio State University emphasizes education, research, and the professional practice and application of computer and information science. The educational program offers undergraduate and graduate degrees through the Ph.D. The research activities which are a central part of the program consist of a broad conceptual base supported by a number of contracts and grants as well as by the university. The broad core research program and these other research tasks interact to form an integrated framework.

Table 1: Growth of Department of Computer and Information Science

	SEPT'67	SEPT'68	SEPT'69	SEPT'70	SEPT'71	SEPT'72	SEPT'73	SEPT'74
A. Staff								
1. Full Time	5	11	14	15	18	18	18	20
2. Part Time	5	9	10	11	12	14	16	12
B. Graduate Students	32	89	114	151	165	187	209	214(est)
C. Undergraduate Students	100	143	300	485	576	450	510	475(est)
D. Course Enrollment (Autumn Quarter)	542	770	1059	1293	1447	1676	1728	1825(est)

	'67-'68	'68-'69	'69-'70	'70-'71	'71-'72	'72-'73	'73-'74	'74-'75
Students Taught	1977	2892	3933	4703	5174	5600	6129	6823(est)
M.S. Degrees Awarded	7	17	35	44	47	49	67	67(est)
Ph.D. Degrees Awarded					4	8	4	15(est)
Applications for Graduate Study	181	190	343	425	400	323	290	
Number of Graduate Students Supported	27	72	78	88	89	83	78	

Undergraduate Programs

Undergraduate degrees in computer and information science are available to students in the College of Engineering, the College of Mathematics and Physical Sciences of the College of the Arts and Sciences, and the College of Administrative Science. The particular program chosen depends upon the student's interests and career objectives.

The undergraduate program in the College of Engineering leads to the degree of Bachelor of Science in Computer and Information Science. This program is designed for the student who wants to specialize in computer and information science from within an engineering environment. Hence, the program provides the student with a core of computer and information science, mathematics, and engineering science. Both depth and breadth in computer and information science, are assured by specific required course sequences in several areas of engineering and science yet, sufficient flexibility exists so that a student can elect a portion of his technical course work in order to develop his individual interests.

There are two undergraduate programs in the College of Mathematics and Physical Sciences. These programs lead either to the degree of Bachelor of Science or the degree of Bachelor of Arts with a major in computer and information science. The programs are cast in a liberal arts setting and are similar in content. The Bachelor of Science program provides a somewhat more technical and thorough education in computer and information science and mathematics while the Bachelor of Arts program is somewhat more flexible and provides an opportunity to relate computer and information science to some other discipline.

The undergraduate program in the College of Administrative Science leads to the degree of Bachelor of Science in Business Administration with a major in computer and information science. This program is designed for the student that is business oriented and desires an education in computer and information science and a general education in the administrative sciences. The program's objective is not to make a computer specialist out of a student, but rather to enable him to recognize the opportunities to use the computer in his managerial activities, to know what to expect from it, and to know how to communicate effectively with computer specialists so that computerized projects will be properly handled from a technical as well as a managerial point of view.

Graduate Programs

The Department of Computer and Information Science offers graduate programs leading to both the Master's and Ph.D. degrees. The graduate program leading to the Master's Degree is available in seven options.

Option I for the student desiring a theoretical foundation in computer and information science.

Option II for the student specializing in information systems.

Option III for the student specializing in computer systems.

Option IV for the student specializing in numerical analysis.

Option V for the student specializing in operations research.

Option VI for the student specializing in biomedical information processing.

Option VII for the student specializing in administrative science.

Each of these options provides a background in several aspects of computer and information science, as well as additional mathematical sophistication appropriate to the student's interest. Each of the options may lead to the Doctoral program in computer and information science, and each may be taken with a thesis option or without a thesis option. (See Appendix A for a listing of courses by number and title.)

All courses of study at the Master's level require completion of a core program in computer and information science, together with the required courses specified for one of the options and additional courses as specified by the student's adviser. The core program includes courses on; Principles of Man-Machine Interaction, Numerical Analysis, Data Structures, Advanced Computer Programming, Digital Computer Organization, Mathematical Froundations of Computer and Information Science, Introduction to Linguistic Analysis, Modern Methods of Information Storage and Retrieval, and Advanced Seminar in Computer and Information Science.

The graduate program leading to the Doctoral Degree in Computer and information Science is flexible in that it is tailored to the particular background and interests of the individual student. These interests may lie in any one of the research and instructional areas already listed as well as in many other cognate areas. A cognate field is defined as a field supporting or closely related to the fourteen Departmental fields and is ordinarily specified by an integrated program of study in other departments of the University.

Course Offerings

Currently there are about 81 courses (each one quarter in length) offered by the Department, 22 of which are largely undergraduate with the remainder being upper level undergraduate and graduate courses. In addition to these courses there are over two hundred courses offered by a variety of departments of the University which are of interest to our graduate students who are encouraged to take these courses.

Faculty

The Department of Computer and Information Science has a full time faculty of twenty members at the assistant professor level and above. They have a wide range of backgrounds and experience. The above faculty is supplemented by staff who have joint appointments with other departments; by staff from other departments who teach courses primarily for Computer and Information Science students; and by adjunct staff people who are employed in off campus organizations who teach courses in the Department of Computer and Information Science (see Appendix B). There are currently a total of about 14 supplemental staff in this category.

INTERACTION WITHIN THE UNIVERSITY

Both the Research Center and the Department of Computer and Information Science interact with other departments and research programs within the University. This is essential because of the multi-disciplinary nature of the activities encompassed in this field. A number of the academic faculty have joint appointments in other departments. Staff members of the Department of Computer and Information Science have appointments in the following departments and organizations:

a. Mathematics
b. Psychology
c. Biophysics
d. Electrical Engineering
e. Accounting
f. Engineering Graphics
g. Instruction and Research Computer Center
h. University Systems Computer Center
i. University Libraries

INTERACTION WITHIN THE INFORMATION SCIENCE COMMUNITY

Columbus, Ohio is one of the major centers for information science and for the transfer of information in the United States. A number of organizations are involved with the activities of computer and information science. This affords an opportunity for students and faculty to interact with appropriate personnel in these organizations. Some of these are:

a. Chemical Abstracts Service
b. Battelle Memorial Institute
c. Bell Laboratories
d. City National Bank
e. Columbus and Southern Ohio Electric Company
f. Western Electric Corporation
g. Rockwell International
h. Industrial Nucleonics Corporation
i. State of Ohio Department of Finance; Department of Highways
j. Highlights for Children
k. Columbus Board of Education
l. Ohio College Library Center

There are a large number of scientists who come to Columbus in order to visit with the Department and Center and who usually present a seminar. (The seminars for the period of this report are listed in Appendix C.) The people cover virtually all phases of computer and information science.

In addition, our people interact at most of the major technical meetings in this country as participants giving papers, assisting on panels, as attendees, and as officials. Hardly a major technical meeting in the appropriate fields is held without a contribution from one or more of the personnel from the Ohio State Computer and Information Science Research Center. A list of these activities can be found in Appendix D.

Research efforts of the staff are disseminated to the professional community through several publication channels. A list of current publications of the Research Center staff is included as Appendix E. In addition, the Research Center issues a technical report series (see Appendix F).

DOCTOR OF PHILOSOPHY DEGREE

The Doctor of Philosophy degree was awarded to the following students during 1973-74.

Name	Dissertation
Paul Sui-Yuen Chan	An Investigation of Symetric Radix for Computer Arithmetic
Mark Lee Gillenson	The Interactive Generation of Facial Images on a CRT Using a Heuristic Strategy
Stephen Philip Hepler	Use of Probabilistic Automata as Models of Human Performance
Paul Tiing Renn Wang	Bandwidth Minimization, Reducibility Decomposition, and Triangulation of Sparse Matrices

RECENT TECHNICAL REPORTS

The Computer and Information Science Research Center began publishing a technical report series in 1968. A list of recent reports of the research center follows (See Appendix F for complete list).

BAUM, R. I.; HSIAO, D. K. A data secure computer architecture. July, 1974. 22p. (OSU-CISRC-TR-73-10)

BIERMANN, A. W.; BAUM, R.; KRISHNASWAMY, R.; PETRY, F. E. Automatic program synthesis reports. October, 1973. 62p. (OSU-CISRC-TR-73-6)

COHEN, D.; LIU, M. T. Emulation of computer networks by microprogrammable microcomputers. June, 1974. 17p. (OSU-CISRC-TR-74-4)

HSIAO, D. K. Logical access control mechanisms in computer systems. July, 1973. 24p. (OSU-CISRC-TR-73-4)

HSIAO, D. K.; KERR, D. S.; MC CAULEY III, E. J. A model for data secure systems (Part I). February, 1974. 39p. (OSU-CISRC-TR-73-8)

HSIAO, D. K.; KERR, D. S.; NEE, C. J. Context protection and consistent control in data base systems (Part I). February, 1974. 28p. (OSU-CISRC-TR-73-9)

KIRSCH, B. M. An improved error diagnostics system for IBM System/360-370 assembler program dumps. June, 1974. 60p. (OSU-CISRC-TR-74-3)

REAMES, C. C.; LIU, M. T. Variable-length message transmission for distributed loop computer networks (Part I). June 1974. 9p. (OSU-CISRC-TR-74-2)

RUSH, J. E.; PEPINSKY, H. B.; LANDRY, B. C.; MEARA, N. M.; STRONG, S. M.; VALLEY, J. A.; YOUNG, C. E. A computer assisted language analysis system. February, 1974. 71p. (OSU-CISRC-TR-74-1)

WANG, P. T. R. Bandwidth minimization, reducibility decomposition, and triangularization of sparse matrices. September 1973. 162p. (OSU-CISRC-TR-73-5)

WYRICK, T.; HSIAO, D. K. System configuration study on the interconnection of the IBM 370/165, DEC System-10 and MICRO 1600/21 computer systems. October, 1973. 26p. (OSU-CISRC-TR-73-7)

II. INFORMATION STORAGE AND RETRIEVAL

ALGORITHMIC CONTROL OF INDEX TERM SPECIFICITY IN DOUBLE-KWIC COORDINATE INDEXING SYSTEMS

The specificity and number of index terms covering a particular area of subject matter in an index is generally a function of the number of articles in that subject area - i.e., the greater the number of articles, the greater the specificity and number of index terms covering those articles in the index. Algorithms for control of index term specificity have been developed and utilized in our most recent version of the Double-KWIC Coordinate Indexing System. The algorithm which has been programmed utilizes statistical data and user specifiable control parameters to select index terms of appropriate specificity from program-generated sets of hierarchically-related index phrases having common words as parent nodes in the hierarchical trees. The current system is being evaluated in conjunction with other fundamental studies on improvement of automatic indexing.

A. E. Petrarca, W. M. Lay, W. S. Stalcup. (Sponsor: National Science Foundation GN 534.1, and GN 27458)

AUTOMATIC INDEXING OF OUTPUTS FROM RETROSPECTIVE SEARCHING OF COMPUTER-READABLE FILES

Interface programs have been written to provide KWIC and Double-KWIC indexes of outputs from retrospective searches of the Ohio State University Mechanized Information Center (MIC) data bases. This capability, in addition to providing a potentially useful service to users of the MIC retrieval system, will be useful for ongoing studies of automatic vocabulary control in printed indexes derived from natural language text and will also be useful for contemplated evaluative studies of the MIC retrieval system.

A. E. Petrarca, J. H. Hsu. (Sponsor: National Science Foundation. GN 27458)

AUTOMATIC INFORMATION SYSTEMS ANALYSIS THROUGH EVALUATION OF SIMULATED PERFORMANCE

The Auxiliary Storage Simulator (AUXSYM) provides execution time commands for between simulation modification. An existing information system model can be modified by (1) substituting one facility for another, (2) equating idependently defined facilities, (3) restructuring the model configuration, and (4) redefining the data base component. AUXSYM also provides, as output, a comprehensive set of utilization and queueing statistics for each information system resource. The current research uses these modification commands and statistics as a basis for fine tuning an original model towards optimum system performance. A definition of optimal, or satisficable optimal, performance, and appropriate measures must be determined. The automatic analysis phase of AUXSYM uses multi-objective mathematical programming techniques to evaluate the output of a simulation against modeler defined performance criteria. Heuristic procedures then decide which modification commands are appropriate to drive the model towards optimality

T. DeLutis, J. S. Chandler. (Sponsor: National Science Foundation. GN 36622)

AUTOMATIC VOCABULARY CONTROL IN PRINTED INDEXES GENERATED FROM NATURAL-LANGUAGE TEXT

Despite the degree of sophistication attained in automatic indexing, much remains to be done in the area of automatic vocabulary control to minimize the amount of concept (or subject) scattering encountered in the indexes so produced. Such scattering is found, for example, in the alphamerically ordered set of keywords mill, millenia, miller, millimeter, million, and mills, where three words related to the concept of "mill" are scattered by intervening, unrelated entries.

Automatic vocabulary control, like manual control, has heretofore been attained through use of a thesaurus (dictionary) appropriate to the subject area covered by the index. However, the thesaurus approach suffers from problems relating to the size and coverage of the dictionaries involved. That is, a thesaurus is often either too small for adequate vocabulary control, or too large to conveniently store, maintain, or even use efficiently. These problems can be significantly reduced by a vocabulary control scheme which is predominantly algorithmic rather than dictionary-driven.

Algorithmic control can be achieved by use of an appropriate stemming algorithm to identify conceptually related words on the basis of a common stem, followed by an appropriate recoding algorithm to construct the preferred index term from the stem and a suitable suffix, if necessary. We have used such an approach to a limited

extent to eliminate scattering caused by singular and plural words in Double-KWIC Coordinate Indexes. A number of stemming-recoding algorithms are presently being investigated for extension of the algorithmic approach to other variant word forms.

A. E. Petrarca, W. S. Stalcup. (Sponsor: National Science Foundation. GN 27458)

DECISION THEORETIC APPROACH TO AUTOMATIC DOCUMENT CLASSIFICATION

Given a set of classes and a set of documents to be classified into these classes, the sequential algorithm obtains keywords for a document and after each keyword computes a new set of aposteriori probabilities of the classes. Classes whose probabilities do not exceed a given threshold are dropped from consideration. This step is repeated until there is either just one class remaining or the entire document has been read. If there is just one class remaining, the document is classified into that class.

This research attempts to develop a mathematical model of the sequential algorithm. Current areas of research are:

1. development of a decision-theoretic tree to model the process;
2. development of analytical criteria for dropping classes at each stage;
3. development of a method which will separate the "noise" keywords from the good ones;
4. development of analytical criteria for classification into multiple classes.

L. J. White, G. Kar (Sponsor: National Science Foundation, Office of Science Information Service. GN 36340)

INFORMATION PROCESSING SYSTEMS SIMULATION

The goal of the research is the development of a methodology for the performance evaluation of information systems. Initially the research is being restricted to the analysis of the auxiliary storage subsystem within the total system, and is an expansion of a methodology for auxiliary storage analysis developed by the principal investigator. The research is divided into two major items, (a) continued verification and extension of the features comprising the current system, and (b) formulation of systematic procedures, based on a set of performance measures, for design modification and acceptance.

Currently, the methodology is a set of general design procedures capable of evaluating auxiliary storage system designs prior to their implementation. It is predicated on the concept of macro-analytic design evaluation, that is, the comparative evaluation of alternate

systems to identify critical design variables. It represents a unique capability because it combines both hardware and software considerations in the evaluation process. Further, it allows for the analysis of stable systems operating in a batch environment to dynamic systems requiring real time operations. Item (a) will improve these capabilities and increase the scope of applicability of the methodology.

The procedures being developed under item (b) are based on the performance measures provided via model evaluation. Design modification will involve changing system variables for the purpose of improving system performance with respect to imposed time and/or cost constraints. Design acceptance will deal with the variances in system behavior under changes in system parameters, particularly the demands placed on the system and the information structure characteristics of the data base.

T. DeLutis. (Sponsor: National Science Foundation. GN 36622)

N-GRAM KEYTERMS IN AUTOMATIC DOCUMENT CLASSIFICATION

An n-gram is a string of n characters which can be used as a keyterm in automatic document classification or retrieval. Blanks may occur within the body of an n-gram; therefore it is possible for single n-gram keyterms to span work boundaries and thus incorporate word dependencies within a phrase. Since n-grams overlap greatly and the too-frequent ones are assigned little weight, n-grams compress storage of keyterms and require no stoplist or stemming algorithm. A simple counting operation on neighboring n-grams can simulate a parsing routine to equate phrases which differ syntactically but are semantically the same. The present research will investigate these advantages and compare the relative effectiveness of n-gram keyterms to keywords over various automatic classification methods and data bases. Evaluation criteria will be storage and time considerations as well as classification accuracy.

S. E. James

PERFORMANCE EVALUATION OF DATA BASE MANAGEMENT SYSTEMS

The data base characterization facilities of the Auxiliary Storage System Simulator (AUXSYM) permit the analyst to view a data base in two ways: (1) in terms of the logical organizations of its data sets, and (2) in terms of its physical storage organization. A data base is characterized by specifying (1) the logical organizations of its data sets, (2) its areas of secondary storage (physical address spaces), and (3) logical address spaces for its data sets, together with a mapping from logical address spaces to physical address spaces.

The logical organization of a data set is described in terms of the file organization and the type of file access. Data base characterization allows the analyst to change the mappings between data set logical address spaces and physical address spaces and the mapping of physical address spaces to mounted volumes at simulation time. These characteriztion facilities are being used in conjunction with current research on the performance evaluation of data base management systems. These facilities are being used in the development of a methodology to evaluate alternative Data Base Management System (DBMS) features. Input to this methodology are DBMS characterizations. Outputs include data set definitions and procedures which are subsequently incorporated in an AUXSYM model of the particular information system being evaluated.

T. DeLutis, J. Aitken. (Sponsor: National Science Foundation. GN 36622)

A QUANTITATIVE MEASURE OF CONCEPT SCATTERING FOR EVALUATION OF AUTOMATIC VOCABULARY CONTROL TECHNIQUES

The purpose of introducing automatic vocabulary control in printed keyword indexes generated by computer from natural-language text is to reduce the amount of concept scattering that would otherwise be found in such keyword indexes. In order to evaluate the effectiveness of any vocabulary control technique in this regard, it is desirable to have an objective and quantitative measure of concept scattering encountered in keyword indexes subjected to varying degrees of vocabulary control.

Since current indexing theory does not provide any guidelines on how to quantify scattering in an index, a scattering measure based on analogies between indexing systems and thermodynamic systems is proposed. Such analogies apply because an index, like matter at given conditions, may be described in terms of the amount of order in its component parts (i.e., concepts in an index; molecules in matter). For example, in a keyword index produced without vocabulary control from natural language text, each entry is sorted (ordered) alphamerically on its keywords. However, a certain amount of concept scattering (i.e., concept disorder) is inevitably also present in such an index. That is, although the words (or phrases) have been ordered, the concepts which they represent may still be partially disordered. This may be seen, for example, in the alphamerically ordered set of keywords ski, skid, skidded, skidding, and skied, where the first and last keywords refer to the same concept, "skiing", but they are separated by three keywords unrelated to that concept.

Extending the thermodynamic analogy further, the concept scattering of an index may be informally defined as concept entropy. Just as work expended by a refrigeration unit reduces entropy present in an insulated system, work done by a vocabulary control process reduces concept entropy present in an index. Furthermore, "absolute zero"

is meaningful in an indexing system because, in direct correspondence to thermodynamics, when no more concept scattering can theoretically be removed from an index, there remains an "ideal index" with zero concept entropy. Thus, the amount of concept scattering present within an index can be computed by determining the extent to which that index departs from its ideal index. The particular equations and units for best quantifying scattering measurements are being investigated along with the development of ideal indexes to be used for these measurements.

A. E. Petrarca, W. S. Stalcup. (Sponsor: National Science Foundation GN 27458)

A SEQUENTIAL APPROACH TO AUTOMATIC DOCUMENT CLASSIFICATION

This research represents a study of a sequential approach to automatic document classification. A sequential method has the advantage of requiring substantially less computation time than other classification techniques, since the entire document to be classified need not be scanned. The classification accuracy appears to be comparable to other automatic classification methods, as 75 to 87% accuracy on the American Institute of Physics' SPIN abstracts has been achieved using the sequential method. The variation of algorithm parameters has been shown to control accuracy and computation time. Automatic keyword selection techniques have been developed based on frequency of occurrence criteria. A clustering algorithm has been used to generate and check classifications for sample documents based on an analytical measurement of the "distance" between two classifications. Cross-comparisons between the discrimination method of John H. Williams and the sequential method are presently being conducted to evaluate the comparative speed, storage requirements, and accuracy of each method. A mathematical sequential model is also under development which will address the problems of multiple class memberships and of a "general class" for unclassifiable documents. A number of data bases have been acquired for further tests which will establish the stability of the sequential method in various environments.

L. J. White, S. E. James, G. P. Jones, G. Kar, D. E. Westbrook (Sponsor: National Science Foundation, Office of Science Information Service. GN 36340)

III. HUMAN INFORMATION PROCESSING

HUMAN EXTRAPOLATION OF STRINGS GENERATED BY ORDERED CYCLIC FINITE STATE GRAMMARS

The investigation of human behavior in relation to the complexity of sequential concepts can be decomposed into the devising of a measure of performance and a measure of complexity for sequential concepts. In this study, learning was measured by errors made in extrapolating character strings. The strings used were ordered by structural parameters of the grammars used to generate the strings. These were the size of the terminal alphabet and the level of embedding (chain length) of the generating grammars.

Two experiments were performed using a computerized implementation of a string extrapolation task. In Experiment 1, it was found that differences in the size of the terminal alphabet and the level of embedding of the generating grammars produced statistically significant differences in the number of trials to last error, and the number of errors made in string extrapolation. In Experiment 2, it was found that variations in the chain length of the generating grammars produced statistically significant differences in the number of errors made in extrapolation, corrected for trials where there was insufficient information to correctly continue the strings. Finally, for all grammars with 12 characters in their terminal alphabets, a correlation of 0.695 was found between the corrected number of errors and the chain length of the generating grammars.

J. L. Beug, R. Ernst

USE OF PROBABILISTIC AUTOMATA AS MODELS OF HUMAN PERFORMANCE

This research studied the use of probabilistic automata as models of human behavior in a particular task. Subjects were shown repeated brief exposures of an item array and then asked to reproduce the array. Exposure duration and number of presentations before reporting were varied and were both found to significantly affect the number of items correctly reported.

Two models were developed that fit the data closely. Two conclusions were drawn from the first of two models. While subjects probably paired items in the stimulus, recall of the stimulus was one item at a time. Also, changes in the number of items subjects were able to report from one condition to the next were studied with the model. In many cases, changes in performance were attributed to a change in the probability of acquiring one particular item.

Two conclusions were also drawn from the second model. The probability estimates indicated some grouping of items for study by subjects. The effect of the number of items previously acquired on the probability of acquiring an additional item was also studied. The results showed the number of items previously acquired to have a strong effect on this probability.

S. P. Hepler, R. Ernst

IV. INFORMATION ANALYSIS

A GENERAL THEORY OF INFORMATION FLOW AND ANALYSIS

A generalized framework for the development of a theory of information flow which permits the analysis and quantification of information has been suggested. In addition to its theoretical and conceptual interest, there are major and immediate implications for the development of information systems and networks as well as for the general understanding of information flow, retrieval, and transfer.

Numerous intuitive notions exist about the interrelationship between information and decision-making. At the pragmatic level, information has value to the extent that it is useful as a resource for purposeful activity. The primary "purposeful activity" in life is decision-making. Hence, information and decision-making are inextricably tied together. In our formulation, in fact, information is defined in just that way, as being data of value of decision-making.

As a consequence, a measure of the amount of information in a data set or message is defined in terms of a quantity called the *decision state of a decision-maker*. The decision state is a function of the determinism of the decision-maker. We suggest a way of evaluating the decision state quantitatively in terms of a specific decision rule.

We assume that the decision-maker selects probabilities of choice for each alternative available that are proportional to the expected value of each alternative. The *value of the decision state* is then defined as the summation of the expected values of all the possible courses of action weighted by the probability of each course of action. A new measure for the *information contained in a particular decision state* is developed:

$$I = m \sum_{i=1}^{m} \{P(a_i)\}^2 - 1,$$

where m is the number of courses of action available to a decision-maker and $P(a_i)$ is the probability of each course of action. The information is defined in terms of a two-choice deterministic situation which we call a "binary choice unit". This measure is universally applicable for all information that is concerned with the effectiveness of the data upon the recipient.

A measure of the *amount of information in a data set or message* can be arrived at by computing the difference in the amount of information in the decision state before and after receipt of the data. That is, the amount of information is arrived at by considering the impact this new data has on the decision-maker's decision state. In symbolic terms, $I(D)$, the amount of information in data set D, is

$$I(D) = I_{t+1} - I_t,$$

where I_{t+1} and I_t are the amounts of information in the decision state after and before receipt of the data set.

The suggested information measure leads to a measure of the pragmatic information content of a data set for a *particular* decision-maker at a *particular* point in time. The data acquired, processed, stored, and disseminated by an information system may be used, however, as a resource by *various* decision-makers at *various* points in time. Hence, in the design and development of information systems, there exists a problem whose level of complexity is an order of magnitude above that of the primary problems addressed in this study -- the problem of quantifying the information contained in a data set in terms of its overall usefulness for a *range of decision-makers over a period of time*.

One possible approach to this problem of assigning a number to a data set to indicate its composite information content would be to start by determining the relationship between the effectiveness of a decision-maker. Since what is really desired is some indication of the information content of this data set for this decision-maker over a period of time, one may determine some index $\overline{I(D)}$ of the average information contained in data set D over some period of time. Then, if one were to develop a measure of the effectiveness of each of the decision-makers for whom this data set serves as a resource, it would be possible to formulate an *information profile* for the data set. Such an information profile would indicate the average information content of a data set as a function of decision-maker-effectiveness.

If such a profile could be determined for every data set to be stored in an information system, then some number derived from this profile could serve as an index of the composite value of this data set. This method would be of major importance for the development of a sound procedure for the design of more effective information systems.

M. C. Yovits, J. G. Abilock (Sponsor: National Science Foundation. GN 31628)

V. LINGUISTIC ANALYSIS

PARALLEL RECOGNITION OF FORMAL LANGUAGES

Cellular Automata and Cellular Logic Systems are studied in this work as parallel recognition devices for formal languages. It is the goal of the project to develop descriptions of the classes of languages recognizable "immediately" (i.e.; in time essentially independent of the length of the input string). The classes are differentiated by the "amount of hardware" or quantity of cellular space required, as a function of string-length.

For a string X of length n in a regular language R, n modules of cellular space are necessary and sufficient. For X of length n in a context-free language C, n^4 modules are believed sufficient. It is not known if n^4 will be necessary. Work continues on this problem.

Further work will consider recognition of context-sensitive and deterministic context-free languages.

J. Rothstein, M. Moshell

ON THE SYNTACTIC STRUCTURES OF UNRESTRICTED GRAMMARS

Formal definitions for the syntactic structures of unrestricted grammars have been developed. The traditional forms for grammar productions give rise to "derivation structures" and "phrase structures", two distinct notions of syntactic structure which become indistinguishable from trees in the context free case. Parallel theories have been developed for both kinds of structure. Using the accepted definition of structural equivalence for unrestricted derivations, it has been proven that two derivations are structurally equivalent if and only if they have the same syntactic structure. Structural equivalence is an equivalence relation over the derivations of a grammar, and each equivalence class contains a unique leftmost derivation. The relationship between derivation structures and phrase structures has been studied and it has been shown that the two concepts are non-isomorphic.

H. W. Buttelmann

GENERALIZED FINITE AUTOMATA ON UNRESTRICTED SYNTACTIC STRUCTURES

The syntactic structures of unrestricted grammars fall into two non-isomorphic classes: "phrase structures" and "derivation structures". A generalization of finite automata to accept these structures has been defined and studied. Corresponding to the Chomsky hierarchy of grammars, there is a hierarchy of acceptors, and for both kinds of structures, the type 2 acceptors are tree automata. For each type in the hierarchy, the sets of structures recognized by the acceptors are just the sets of projections of the structures of the grammars, and the languages of the acceptors are just the languages of the grammars. Finally, the set of derivation structures of a recursively enumerable language is recursive.

H. W. Buttelmann

A FORMAL THEORY OF THE SYNTAX, SEMANTICS AND TRANSLATION OF PHRASE-STRUCTURE LANGUAGES

A formal definition for a semantics for phrase structure grammars, called a phrase structure semantics, has been developed. It is a model of the following semantic philosophy: 1) it is phrases which have meaning, and 2) the meaning of a phrase is a function of its syntactic structure and of the meanings of its constituents. A pair (G,S) where G is a phrase structure grammar and S is a phrase structure semantics, is called a phrase structure language description. The languages of psld's are just the recursively enumerable sets. It has been shown that for any psld with a type 0 or type 1 grammar, there exists a psld with a context free grammar that defines the same language, with the same meanings. Translation is defined on the languages of psld's, and it has been shown that the translation function on the languages of arbitrary pairs of psld's is effectively computable.

H. W. Buttelmann

SYNTAX-DIRECTED AND SEMANTIC-DIRECTED TRANSLATION OF PHRASE STRUCTURE LANGUAGES

The basis of this theory is a model, from ealier work, of language description called a phrase structure language description, which contains both syntax and semantic information. Certain translation algorithms for these languages are being studied, which promise to be more efficient than the general algorithm of earlier theory. These algorithms are either syntax-directed or semantic-directed, depending on whether they are controlled by syntax information exclusively, or by both syntax and semantic information. A particularly fast translation algorithm is a special case of the syntax-directed translations, where the syntax control information can be specified in a simple, finite way. It has been shown that it is possible to use a semantic-directed translation scheme as a translator generator, to produce the finite specifications of a very fast syntax-directed translation, if such a translation exists.

Running computer programs of the translator and translator-generator have been developed, and the programs are being used to develop and test definitions and translations of very simple languages.

H. W. Buttelmann

SEMANTIC-DIRECTED TRANSLATION--ANOTHER APPROACH

Let L be a language and U a universe of discourse. A semantics for L is a specification of a function $f: L \to 2^U$, which assigns to each sentence w of L one or more meanings f(w) in U. We are investigating methods for specifying the semantics of context free languages so that a partial function $f^{-1}: 2^U \to L$ exists and is effectively computable under certain conditions. Then a translation τ from L_1 to L_2 can be effectively computed as $\tau(w) = f_2^{-1}(f_1(w))$.

H. W. Buttelmann, A. Pyster

STUDIES IN DECIDABILITY, EXISTENCE, AND EFFECTIVENESS OF TRANSLATIONS ON PHRASE-STRUCTURE LANGUAGES

A theoretical model for the semantics of phrase structure languages is taken from earlier research. The phrase structures of sentences are used to compute semantic functions. The model is applied to a theory of translation between phrase structure languages. Issues of the decidability of the existence of a translation and effectiveness of a translation are studied. Various types of translation are described under different constraints, and decidability and effective computability issues are examined for each type. Necessary and sufficient conditions for existence for various types of translation are being studied.

H. W. Buttelmann, R. Krishnaswamy

SOME PROPERTIES OF SYNTAX-DIRECTED TRANSLATIONS

We are investigating problems relating to the existence and computability of finitely specified syntax-directed translations. The properties of the set of syntax trees generated by a finite generating set of trees, and maps on those sets of trees, are being studied, and we are investigating the following conjecture: Let L_a be a phrase-structure language of numbers written in radix a and L_b a phrase-structure language of numbers written in radix b. Then a finitely specified syntax-directed translation from L_a to L_b exists if $\log_a b$ is rational.

H. W. Buttelmann, F. J. Dickey

THE ROLE OF THE NONTERMINAL IN LANGUAGE GENERATION

We examine in depth the differences between grammars which allow nonterminals and those which do not. We do so by introducing a terminal-only grammar, which is a rewriting system in which nonterminals in their usual capacity and status do not exist.

A hierarchy of grammars and languages, parallel to that of Chomsky's, is formed and investigated. The properties of the terminal-only hierarchy are markedly different from those of the Chomsky hierarchy and many results on these differences are proven. In addition, we discuss some of the functions of nonterminals and why not having them in terminal-only grammars drastically affects the generated languages.

H. W. Buttelmann, A. Pyster

VI. ARTIFICIAL INTELLIGENCE

DECISION THEORETIC MODELS FOR INVARIANT PATTERN RECOGNITION

The problem of statistical classifier design when pattern vectors depend upon variables, such as orientation, external to the feature space is investigated. The problem arises in designing an automatic identification system for aircraft using low frequency radar data. In such a situation it is desirable for the feature selection and decision algorithms to be either independent of the external variables or to utilize any available information regarding them. A niminum error rate has been obtained using discriminant functions based on statistical averaging over the external variables. Efficient methods have been developed for computing Bayes discriminant functions when class-conditional probability density functions are either mixtures of multivariate normal or Rician density functions. Treating each discrete orientation as a sub-class, an efficient sequential technique based on the generalized sequential probability ratio test has been obtained and tested in Monte-Carlo simulations. The performance of the sequential classifier has been compared to the theoretical performance of a classifier based on majority voting over individual decisions. Other sequential and non-parametric techniques, and classifier design when samples correspond to different orientations are currently being investigated.

L. J. White, S. N. Srihari (Sponsor: Air Force Office of Science Research Grant 69-1710)

DECISION THEORY FOR PATTERN RECOGNITION

If $p(x)$ and $q(x)$ are the densities for the N-dimensional measurement vector x, conditioned on the classes c_1 and c_2, and if finite sets of samples from the two classes are available, then a decision function based on estimates $p(x)$ and $q(x)$ can be used to classify future observations. In general, however, when the measurement complexity (the dimensionality, N) is increased arbitrarily and the sets of training samples remain finite, a "peaking phenomenon" of the following kind is observed: classification accuracy improves at first, peaks at a finite value of N, called the optimum measurement complexity, and starts deteriorating thereafter. In this research, we have derived, for the case of statistically independent measurements, general conditions

under which it can be guaranteed that the peaking phenomenon will not occur, and the correct classification probability will keep increasing to value unity as $N \to \infty$. Several applications are considered which together indicate that, contrary to general belief, independence of measurements alone does not guarantee the absence of the peaking phenomenon.

B. Chandrasekaran, A. K. Jain (Sponsor: Air Force Office of Scientific Research. 722351)

DISTANCE FUNCTIONS FOR INDEPENDENT MEASUREMENTS AND FINITE SAMPLE SIZE

It is generally known that the "further apart" one makes the class-conditional densities, the smaller will be the probability of misclassification determined. The distance functions have also been found to be quite useful to the recognition system designer in selecting the best subset of features from a given set. Most of the distance functions are applicable only when the class-conditional densities are known; how these will behave if the estimated densities are substituted for the true ones is generally uncertain. In this research, a distance function has been derived when the features are independent and only a finite number of samples is available to estimate the class-conditional densities.

B. Chandrasekaran, A. K. Jain (Sponsor: Air Force Office of Scientific Research. 722351)

FINITE MEMORY DECISION THEORY

In this research we investigate finite memory, deterministic algorithms for the symmetric hypothesis testing problem. It is shown how members of a class of irreducible, deterministic, finite memory algorithms can be constructed to give a steady state probability of error that decreases asympathically faster in the number of states than the best previously known deterministic algorithm. It is strongly believed that the optimal algorithm lies in this class.

B. Chandrasekaran, C. C. Lam

GRAMMATICAL INFERENCE PROBLEM

The problem of grammatical inference for finite state grammars can be approached analytically, and has application in syntactic methods for pattern recognition. The current project is implementation of a technique proposed by Pao in 1969 which, given a set of strings from the target language L, will first construct a machine M_0 to accept exactly those strings. Next the algorithm constructs a lattice of all those machines which can be produced from M_0 by merging its states.

The algorithm then systematically searches the lattice to find the target machine by examining pairs of machines M_1 and M_2. Using a string x accepted by exactly one of M_1, M_2 and an oracle to determine whether $x \in L$, one of M_1, M_2 is eliminated from consideration and the other maintained in the lattice. Work has been done on the storage of this latice and on the construction of a minimum M_0. Currently the problem of defining an efficient "systematic" search of the lattice is being examined as well as the problem of finding a string x for the test between M_1 and M_2. Future research plans are to extend the solution to context-free grammars, and to apply these results to pattern recognition problems.

L. J. White, D. A. Marik (Sponsor: Air Force Office of Science Research. Grant 72-2351)

INTERACTIVE GENERATION OF A HUMAN FACIAL IMAGE ON A CRT USING A HEURISTIC STRATEGY

Sketching a human face is a task which involves spatial decisions and a knowledge of aspects of the face that are important in recognition. These are talents which non-artists lack. A man-machine system has been developed with which a non-artist can create, on a graphic display, any male caucasian facial image from a photograph in front of him. The computer system contains pre-stored facial features, an average face used as a starting point, and a heuristic strategy which provides the user with the "how-to-do-it" knowledge necessary. The user makes all of the visual decisions and can make changes in individual features or hierarchically determined sets of features through analog input devices.

B. Chandrasekaran, M. L. Gillenson (Sponsor: National Science Foundation. GN 534.1 and GJ 41422X)

AN INTERACTIVE SYSTEM FOR AUTOMATIC PROGRAM SYNTHESIS

A man-machine interactive computer programming system has been developed which enables a person to work through example calculations using a light pen at a CRT and which then automatically synthesizes a program for doing those calculations. The system enables the user to "show" a computer how to do a calculation rather than having to write a program in the traditional manner. This system, called an autoprogrammer, works by storing the complete sequence of actions executed during an example computation and then synthesizing the smallest computer program which is capable of executing the example.

In order to reduce the synthesis time, the program synthesis algorithm utilizes a number of techniques including preprocessing the trace information to reduce enumeration and pruning the search using a failure memory. These have proven sufficiently powerful

to do complex synthesis problems, for example, a universal Turing machine, in less than a minute. The autoprogrammer system has been successfully used to produce many programs from interactive calculations including a finite state machine minimizer and a small compiler. A ten minute color movie has been made to demonstrate the operation of the autoprogrammer system. Work is currently in progress to explore the problem of synthesis from weaker input information, in particular from traces of memory snapshots.

A. W. Biermann, R. Krishnaswamy, F. Petry, B. Chandrasekaran. (Sponsor: National Science Foundation. GN 534.1 and GJ 34739X)

LOW-FREQUENCY RADAR AIRCRAFT DETECTION AND CLASSIFICATION

Previous research has demonstrated the feasibility of using multiple low-frequency radar returns for target classification. Simple object shapes have been successfully classified by linear techniques, but aircraft data poses greater difficulty, as in general such data are not linearly separable. Since two parameters specify the aspect angle, this data lies on a 2-dimensional surface in n-space, where n is the number of frequencies utilized. A bilinear approximation of this surface has provided interpolation capability, improved proximity information, and an intersection algorithm which determines whether the aircraft data is separable. Improved separability together with decreased susceptibility to additive noise was obtained by mixing horizontal and vertical polarization data at an optimum polarization angle. Studies have indicated the importance of a phenomenon known as bias, where in the identification of a test signal as one of two given aircraft, the probability of misclassification for one aircraft can be substantially larger for one aircraft than the other for high noise levels. Thus classification algorithms are currently being developed which jointly minimize the probability of misclassification and the bias effect. Another approach being investigated is multipoint classification, where the input is a sequence of independent radar measurements of the aircraft target to be identified. It has been shown that this approach can substantially reduce the effects of noise even for a modest sample size, and a number of classifiers are being investigated for this purpose.

L. J. White (Sponsor: Air Force Office of Science Research. Grant 69-1710)

NUMBER TREES, SEMIGROUPS, AND FORMAL LANGUAGES

Any function f with integer values, defined on all the integers, where the value does not exceed the argument, defines a set of trees whose nodes correspond biuniquely to the integers. The trees are digraphs whose paths lead to "roots", whose integer labels are given by $f(n) = n$ (strictly speaking the tree part of the graph is obtained by suppressing self-loops, which occur only for such n's).

A set of such functions defines a semigroup of transformations on the integers, with composition of functions giving trees formed by concatenating "tree paths" from sub-trees associated with individual functions. An interesting function, whose complexities seem inexhaustible, is provided by Euler's totient function, $\varphi(n)$. All odd integers are "leaves" in the sense that there is no n such that $\varphi(n)$ is an odd integer > 1. If p is a prime and 2p+1 is not prime then 2p is a leaf. The "main stem" or trunk of the φ-tree consists of powers of 2, ($\varphi(2^n) = 2^{n-1}$) in the sense that the interated φ-function carries all integers into the trunk (example: 2^m3^n joins the trunk at 2^m after n iterations). The number of primitive roots of any prime p are given by $\varphi(\varphi(p))$. Formal languages can be encoded as sets of integers (e.g. by the m-adic or other notations), and generative grammars, Post systems, and Markov algorithms can all be viewed as "growing a number tree from a root", the root being the integer associated with the start symbol, axiom, or input string respectively. The number-tree functions {f}, defined "inversely" by the production or rewriting rules, are then "parsing" operations, and a tree-path to the root is equivalent to a derivation, proof, or computation in the corresponding formal system or language. As numbers in these encodings are vastly smaller than Goedel numbers, they may be useful in mechanizing discoveries of proofs in artificial intelligence, in decision problems, or in parsing problems.

J. Rothstein

VII. INFORMATION PROCESSES IN PHYSICAL, BIOLOGICAL AND SOCIAL SYSTEMS

ON THE ROLE OF POST HOC ERGO PROPTER HOC IN SCIENTIFIC METHODOLOGY AND ARTIFICIAL INTELLIGENCE

The literatures of pattern recognition, artificial intelligence and scientific methodology give so much prominence to decision theory, statistical or logical, that the need for generating hypotheses from which the choices are to be made is often obscured. Work so far on trying to discover the mysterious wellsprings of creativity whence come hypotheses, in the sense of seeing how close one can get to a formal discovery procedure (probably impossible to do, but this does not mean that one can't program much of the drudgery), has led to the following. The "universe of interest" has a "history", a describable aspect of which is to be singled out as an "initial condition", "cause", "specification", "state" or the like, some other aspect being the "later condition", "effect", "subsequent state" etc. What is sought is a "dynamical law" or the like permitting one to deduce the second class of aspects from the first, thereby reducing the (empirical) information needed to write a history. Post hoc ergo propter hoc (usually dismissed with contempt as a logical fallacy) can be understood as a generator of a set of hypotheses which must undergo further winnowing (e.g.; by Bayesian inference from new experiments). It then becomes a useful tool (instead of a snare and a delusion) for iterated experimentation. Creativity comes in at the stage of designing new experiments.

J. Rothstein

VIII. MATHEMATICAL TECHNIQUES

AN ALGORITHM FOR SUB-OPTIMAL SET COVERS WITH CONFIDENCE MEASURES OF OPTIMALITY

An effective heuristic has been developed for the minimum set cover problem: min cx such that $Ax \geq 1$, $x = 0$ or 1, A binary. Two algorithms have been developed: one for the general costs problem (arbitrary c vector), and one for the cardinality problem (all c_i equal). Both algorithms are implemented in computer programs in FORTRAN and assembler. Test results for the cardinality problem exhibit very good performance on test problems from the literature. Current work includes analytic developments for measures of confidence of obtaining a global optimum solution, refinements of the heuristic, and testing the general costs program.

L. J. White, M. E. Doherty

BANDWIDTH REDUCTION OF SPARSE MATRICES BY ROW AND COLUMN PERMUTATIONS

It has been shown that there is an important class of sparse matrices, called well r-bandable, for which bandwidth minimization can be effected simply and deterministically. This analysis has led to an algorithm, called SYMBAND, which determines row and column permutations to achieve a significant bandwidth reduction. The concept of stability of a bandwidth reduction algorithm has been introduced. Extensive computational results have been obtained involving random symmetric sparse matrices, and the algorithm SYMBAND has been shown to be stable, convergent, and reasonably fast.

D. S. Kerr, L. J. White, P. T. R. Wang

BINARY STRINGS AND TOPOLOGY

A continuing interest in the problem of representing topological properties in binary language has led to a procedure for generating strings over (0,1) with topological interpretations. It appears applicable to topological spaces admitting both a measure and a metric. It is convenient to take both as bounded, with the strings associated with it interpretable as binary decimals i.e.; real positive numbers

between 0 and 1. The terminating decimals represent a countable basis of neighborhoods, the infinite decimals represent points of the space, and the chain of neighborhoods represented by arranging all the n-digit strings in numerical order constitutes the n^{th} approximation to a space-filling curve. It is possible to define a metric equivalent to that originally given for which this n^{th} approximation has length $(1-2^{-n})$. The space-filling curve then has unit length, can be used to "coordinatize" the entire space 1-dimensionally, and thereby to induce a natural 1-dimensional metric. Obviously there are neighboring points, according to the original metric, which are separated by finite distance according to the 1-D metric, but this does not detract from the usefulness of the coordinatization to name the points of the space. The non-1-D nature of the space is taken into account by specifying closure or convergence relations in the set of infinite binary strings. While every infinite binary sequence defines a unique point, and points correspond to unique sequences almost always, there is an everywhere dense countable subset of points defined by more than one sequence, analogous to the 1-D identity in the limit of .01000... and .00111.... The number of sequences to be so identified depends on dimensionality. The results suggest a possible new approach to dimension theory as well as the possibility of existence of topological spaces with a generalized dimensionality associated with a set of integers rather than a single integer. The reader is invited to experiment with the following highly illuminating case. The space is a 30^{o}-60^{o}-90^{o} triangle in the Euclidean plane with coordinatization assigned by iterated dissections using altitudes dropped from right angle to hypothenuse, the smaller triangle being 0, the larger 1. As in the Heine-Borel theorem, every infinite binary string defines a unique point. Verify the properties earlier discussed.

J. Rothstein

COMPARISON OF ROUTINES FOR SOLVING NONLINEAR SYSTEMS OF EQUATIONS

The object of this research is to parameterize nonlinear systems of equations so that one can more effectively select which routine is best suited for solving a given nonlinear system. Parameters under consideration include the size of the system, the goodness of the initial approximation to the solution, the nonlinearity of the system, and the number and interplay of the relative extrema of the component functions of the system. Based on the above criteria, a model has been developed which will provide the basis for comparing production routines for solving nonlinear systems of equations and evaluating their effectiveness with respect to systems based on the above parameters.

D. S. Kerr, L. J. White, D. L. Kalmey

MINIMUM-REDUNDANCY VARIABLE LENGTH CODING

In 1952, Huffman provided an elegant algorithm for coding non-equiprobable messages using alphabets of equal cost letters. There is also a well-known efficient algorithm, called the extension technique, for coding equiprobable messages using binary alphabets of unequal cost letters. To date we are not aware of the existence of a simple algorithm for the general case of coding non-equiprobable messages using unequal cost coding letters. Nevertheless, the problem has been attacked by a number of people, including Karp who provided an integer programming solution, and Varn, who obtained a selective enumeration technique.

Our research is aimed at unifying the simple techniques for equiprobable and equal cost cases of this general problem. As a first step we have obtained an algorithm for coding equiprobable messages using two or more unequal cost coding letters. Unlike Varn's algorithm, this algorithm is a straightforward generalization of the extension technique. We have also obtained a few preliminary results that may lead to a better characterization of the problem, and ultimately a reasonably simple algorithm for the general case.

L. J. White, V. Santhanam

SIEVES AND CONTEXT SENSITIVE LANGUAGE

It has been shown that the well-known procedure known as the sieve of Eratosthenes, and a class of generalized sieves, can be implemented on a linear bounded automaton. This leads to simple proofs that not only do the prime numbers, in any base representation, constitute a context sensitive language, but many other languages as well. Examples are the "lucky" numbers studied by Ulam and collaborators, the self-avoiding random walks on general lattices, and many others. A similar analysis proves context sensitivity for numbers satisfying arbitrary systems of linear recursion equations (e.g., the Fibonacci numbers). Though the procedures considered are certainly included both in those it is intuitive to regard as sieves and those which correspond to context sensitive languages, it is not clear whether the two concepts are coextensive, one is a special case of the other, or whether they merely overlap, and if the latter, what kinds of non-CS recursive languages might be recognized by sieves.

J. Rothstein

SOME L^p-NORM MAXMIN PROBLEMS: ALGORITHMS AND APPLICATIONS

Two closely related maxmin problems are under investigation, one involving the Euclidean norm, and the other a rectilinear norm. In each case a point is to be located in a given convex polyhedron to maximize the minimum distance from a given set of points in the

polyhedron. The rectilinear problem is shown to be a linear programming problem with m x n constraints where m is the number of points and n is the dimensionality of the space. The Euclidean problem is shown to be a nonconvex programming problem, specifically one for which a finite candidate solution set can be generated. It has been demonstrated that the well known duality approaches cannot be used to obtain an efficient algorithm for the general n-dimensional problem. However, an efficient algorithm is proposed for the Euclidean problem in the plane. The case where the convex polyhedron is also the convex hull of the given points is considered. The potential relevance of the problem to location theory and pattern classification is under study. An extension of the algorithms to the case of weighted norms is also under investigation.

L. J. White, B. Dasarathy

A STUDY OF THE OPTIMAL PLACEMENT OF KNOTS FOR CUBIC SPLINE APPROXIMATIONS

If the number and position of knots are specified, an optimal approximation of a given function by a cubic spline can be effected either in the sense of mean square error or the Chebychen uniform norm. This research has investigated the optimum position of a given number of knots, and the variation of the above error criteria with the number of knots. This research was facilitated by a computer program developed by C. DeBoor and J. Rice, which iteratively obtains an approximation of the optimal knot locations in order to minimize the mean square error for a given function. One interesting result obtained is that for a large class of functions, as the number of knots is increased, there is a marked decrease in error as a critical number of knots is approached, rather than a gradual decrease as might be expected. This was true even for quadratic and cubic polynomial functions tested, where error rates were quite low, as would be expected for a cubic spline approximation. It was found that the extrema of the second derivative provide a good guide for the placement of knots, and the curvature of the second derivative gives an indication of how many knots should be utilized. This curvature analysis also indicates the critical number of knots for which a substantial drop in the error occurs, and beyond which negligible decrease in error is obtained. It is hoped these studies will yield simple intuitive guidelines for a good choice of knot locations for cubic spline approximations of a given function.

L. J. White, S. Muftic, P. Wong

TOWARD AN ARITHMETIC FOR CELLULAR AUTOMATA AND PARALLEL COMPUTATION

Earlier work on patterns generated by groupoid string operations led to the idea of trying to construct a number system without carrying, based on the cyclic group of order 2. It has been shown that requiring

the number system to be a Gray Code (which is implied by the no-carry property), with the daughter string representing a number double that represented by the original string (considered as a finite non-zero portion of a doubly infinite string) determines the complete number system (up to a mirror reflection, i.e., a convention on direction of reading). Many interesting properties of this encoding of the integers have been discovered. All non-integral numbers are represented by infinite strings, but only those corresponding to terminating binary decimals can yet be evaluated. Uniqueness has also been proved for the ternary system based on the cyclic group of order three. For higher bases it is not known whether uniqueness up to isomorphism is provable.

J. Rothstein

TURING UNIVERSALITY AND PARALLELISM OF GROUPOID STRING ALGORITHMS

Earlier work on patterns and algorithms made use of strings over groupoid alphabets, where a "daughter string" is generated from a given string, letter by letter, using the groupoid algorithm on contiguous pairs of letters of the given string. Using a one-dimensional cellular automaton permits parallel computation of a daughter string in a single step. Explicit constructions have been given to specify a groupoid algorithm equivalent to an arbitrary one-dimensional (1-D) cellular automaton (C.A.) and to specify a 1-D C.A. equivalent to an arbitrary Turing machine. Successive Turing machine configurations are represented by successive generations of daughter strings. This Turing universality of groupoid algorithms has been shown to be realizable with a simple modification of an infinite shift register, namely, where a shifted copy of the entire string is combined with an unshifted version letter by letter (by the groupoid algorithm) in the register cells. This computation is completely parallel in nature and is accomplished in a single step. A one-step parallel procedure for finding a "parent" string (or showing that none exists) has also been given, as has a "speed-up theorem". The latter constructs, from any groupoid, another groupoid in which the k^{th} descendant of the string over the original groupoid is computed in parallel in one step.

J. Rothstein

IX. SYSTEMS PROGRAMMING

CONTEXT PROTECTION AND CONSISTENT CONTROL IN DATA SECURE SYSTEMS

In data secure systems, the basic unit of information for protection is called a data unit, which may be a single data item or a collection of data items having the same properties with respect to access control.

Context protection means that the same data unit may be protected differently in different contexts. For example, the same data field may be protected differently in two different records. The difference may be determined by the manner in which the fields and records are being accessed.

Consistent control is concerned with the problem that when new data units based on the old data units of the data base are created by the users, these data units may be protected consistently in the sense that their access requirements must be generated automatically and be conformed with the access requirements of the old data units.

Both context protection and consistent control can be enforced by means of certain built-in relations among the data units involved. These relations can reveal any violation of context protection and consistent control. A graph-theoretic approach is being used for this study.

D. K. Hsiao, C. J. Nee, D. Wong (Sponsor: Office of Naval Research. N00014-67-A-0232-0022)

A CROSS-ASSEMBLER FOR THE MICRODATA 1600/21 COMPUTER SYSTEM

The objective of this project is to develop an assembler for the Microdata 1600/21 written in PL/I on the IBM 370/165. The assembler will accept an assembly program in card images and produce a source program listing and a hexadecimal object program deck.

A preprocesser will be included that will support conditional assembly and a macro language processing.

D. K. Hsiao, T. Rodeheffer (Sponsor: The Ohio State University. Instructional and Research Computer Center and Department of Computer and Information Science)

CRYPTOGRAPHIC TECHNIQUES AND DATA SECURITY IN DATA BASE SYSTEMS

The importance of data security is well recognized in and out of computer science. Current techniques, such as "access control" are inadequate for protecting against all threats to privacy which exist in computer systems. As Peterson and Turn pointed out, the only safeguard against some of these threats is "cryptographic transformations".

Cryptography has been used extensively in communication systems but its implementation in computer systems require different considerations. The main ones are: the long "eye" of the data and the necessity to process the data in the computer system. This leads to the idea of "processable" ciphers which allow processing of the data in its ciphered form. The first goal of our research is to find such ciphers, which provide reasonable security on the one hand and processing capability on the other hand.

Since cryptographic transformations affect mainly the physical representation of the data, the subject is strongly connected to the problem of organizing and describing data bases. The problems of implementing cryptography as part of current data definition languages will be shown. We will also show the ways to expand these DDL's to include cryptographic transformations. Our second goal of the research is to find how to apply cryptographic transformations effeciently and how they affect the organization and description of data bases. In looking at the problem of cryptography and data bases we found that several other problems, related to security, exist in current description of data base systems. One of them is the centralization of all security information in only one part of the data base, which may cause a threat to its security. Another problem is that security checks in general and cryptographic transformations in particular cause a significant overhead and we are concerned about the efficiency of such measures. The third goal of our research is to build a general model of a data base system which incorporates security considerations into it and which can demonstrate the tradeoff between security requirements and processing efficiency.

E. Gudes, H. S. Koch, F. Stahl

A DATA BASE MANAGEMENT SYSTEM WITH GENERALIZED DATA STRUCTURE, ACCESS AND UPDATE

In a contemporary data base management system (DBMS) only a restricted number of record organizations, file structures, access and update techniques are supported. Of these that are supported, each is considered as a special case, with completely different software routines being used. The purpose of this research is to design and implement an experimental data base management system for generalized record organization and file structure and for unified access and update.

With this DBMS we can then study the efficiency and effectiveness of data base management systems issues and prepare the ground work for the study of data security and integrity problems in data base management systems.

H. S. Koch, B. J. Kammerer, R. Knablein, H. Sieberth (Sponsor: Office of Naval Research. N00014-67-A-0232-0022)

DATA SECURE SYSTEM ARCHITECTURE

The overall goal of this research is the development of a computer (hardware) architecture which facilitates the implementation of data base management systems with logical access control capabilities. In the first phase of the work we investigated the implementation feasibility of different types of data base protection. A data base model consisting of objects, relations among objects and relations among these relations was used to characterize different protection strategies. Two basic protection strategies evolved: concealment protection and information-theoretic protection. A concealment-oriented mechanism protects objects and relations by selectively preventing user access to certain objects. The restrictiveness of a protection mechanism may be characterized as the number on nonexplicitly protected items that must also be protected to insure that the specified security conditions are maintained. Various classes of concealment protection were examined with respect to their implementation complexity and degree of restrictiveness. Concealment protection may be divided into two broad categories: static protection and dynamic protection. Dynamic protection unlike static protection requires an access history. The various degrees of dynamic protection are characterized by the amount of history information that must be maintained and the way derivable relations are used.

The simplest form of dynamic protection prevents effective access to a derivable relation by preventing access to those objects which characterize the protected relation. A more complex form of dynamic protection prevents access to objects and relations which may be used to derive the protected relations. Information-theoretic protection is based on a model of the inference capabilities of the user. Such a protection mechanism dynamically determines the probability that relations are known to the user. This probability is determined by considering all queries asked by the user, all specified security conditions and all inferences assumed available to the user. Protection is provided by specifying a threshold probability on each relation which the system may not allow the user to cross. Information-theoretic protection is by far the most complex form of protection, but potentially provides the least restrictiveness. The next stage of the work will select those protection strategies that are mostfeasible and design a computer architecture which provides them the context of a data base management facility.

D. K. Hsiao, R. Baum (Sponsor: Office of Naval Research N00014-67-A-0232-0022)

A DATA SECURE SYSTEMS MODEL

A multi-level model is developed to allow the different problems in data security to be considered at a level of abstraction appropriate to the specific issue under study. The highest level is conceptual. In it, "patterns of protection" (intuitively, the ways in which the users may access the data) can be defined in formal and unambiguous ways. The intermediate level of the model is structural. Here, the primatives and structures needed to realize the protection patterns of the conceptual level are defined. The role of a protection system is to prevent the completion of illegal access attempts. This can be thought of as being done by the creation of security deadlocks between offending users and "pseudo-users". The common capability list and access list protection mechanisms may thus be viewed as simply different techniques to detect these security deadlocks. Another aspect of the structural model is the definition of logical data base structures which are more amenable to the requirements of data secure systems. On the lowest level of the model, a system to illustrate the utility and practicality of the other, higher levels is being created.

D. K. Hsiao, E. J. McCauley (Sponsor: Office of Naval Research. N00014-67-A-0232-0022)

DESIGN AND IMPLEMENTATION OF A DATA BASE PROTECTION MECHANISM

The goal is to design and implement a protection mechanism for a data base system. In this system, protection violations can be classified into two types, intentional and nonintentional. An intentional violation is an explicit request by the user for data which are denied by the protection pattern. In this case the user is deadlocked by the mechanism. In a nonintentional violation the user is allowed to request data, some of which may be denied by the protection pattern. In this case the user is not deadlocked. He will recieve those data which are not denied.

The design is aimed to show that the working of the mechanism in the case of intentional violations can be characterized by the theory of deadlock and in the case of nonintentional violations by the practice in authority item and capability list systems.

D. K. Hsiao, N. Kaffen (Sponsor: Office of Naval Research. N00014-67-A-0232-0022)

DESIGN AND IMPLEMENTATION OF A METHOD FOR CONTROLLING COMPUTER PROGRAM-POSTMORTEM DIAGNOSTIC OUTPUT

Debugging, or the correcting of programming errors, is one of the major problems in implementing a large computer software system. This project is to design and implement a postmortem debugging aid

which is controlled by the program prior to failure and outputs the most relevant information upon termination. The central part of this proposed debugging aid is the program control block. This block will contain fields of information useful for debugging. The objective of this project is to design and implement a postmortem debugging aid controllable by the program being tested; and to show that the use of this software tool significantly increases the debuggability of common assembly language programs.

R. F. Mathis, M. S. Friedman

THE DEVELOPMENT OF GUIDELINES FOR DATA BASE REORGANIZATION

The major objective is the development of general guidelines and procedures that will allow the need for data base reorganization to be recognized and that will assist in the reorganization process. Specifically, the objectives are to:

1. Examine, identify and classify the different types of data processing installations using Data Base Management Systems (DBMSs).

2. Identify and classify system limitations and Data Base Administrator (DBA) and user requirements with respect to data base reorganization giving the DBA the framework and a set of guidelines with which to work.

3. Identify the types of commands that need to be in the Data Definition Language (DDL) for data base reorganization and specify what needs to be added to the Data Manager to analyze and identify the need for data base reorganization.

4. Identify and develop portions of the Data Manager that will give DBAs more flexibility with the DBMS they have chosen.

H. S. Koch (Sponsor: The Ohio State University. Graduate School)

FILE PARTITIONING AND RECORD PLACEMENT IN ATTRIBUTE-BASED FILE ORGANIZATIONS

In this research a method is suggested which provides for improved accessing efficiencies by recognizing the desirability of logical file partitioning. With such a technique the number of different secondary storage cells (e.g., pages in virtual memory systems) accessed is minimized. File partitioning dictates the record placement in the attribute based file organizations. It is the aim of this research to relate the results to the logical file partitioning problem in the CODASYL DDLC and to the enhancement of protection and access control in shared data base systems.

D. K. Hsiao, F. Manola, E. J. McCauley (Sponsor: Naval Research Laboratory and Office of Naval Research N00014-67-A-0232-0022)

GRAPHICAL MODEL OF FUNCTIONAL DEPENDENCIES IN A RELATIONAL DATA BASE

A relational model for data base management systems has been proposed to allow the user to express logical relationships among the data without having to be concerned with design and implementation details such as storage structures. In this model information is represented as sets of relations. Various normal forms of relations have been suggested in order to make the sets of relations easier to understand and control, simpler to operate upon and more informative to the casual user. When there are many relations in a data base, there are many possible choices of normalized relations. Functional dependencies among the various attributes affect the choice of relations. A graphical model is being developed to express these dependencies so that the best set of relations may be selected.

D. S. Kerr (Sponsor: Office of Naval Research. N00014-67-A-0232-0022)

IMPROVED ERROR DIAGNOSTICS FOR IBM SYSTEM /370 OS/MVT ASSEMBLER PROGRAM DUMPS

A major problem with assembler program dumps as they now exist is that they leave the inexperienced (and sometimes experienced) programmer in the dark as to the actual cause for the program's abnormal termination. In many cases the error cause is only loosely connected to the system message produced with the dump, and the programmer has no alternative but to turn to a consultant for debugging help. To remedy this situation, a postmortem debugging system has been developed to provide the programmer with improved error diagnostics and debugging hints upon abnormal termination of his program. Contents of relevant registers and storage locations are analyzed, along with the program's object code and system control blocks, to arrive at the most probable cause(s) of the exception. The programmer is not deluged with a great deal of what he considers to be useless output. The system currently handles program interruptions (types OCX), but will be expanded to cover many other error types. Additional diagnostic features, including selective facilities, are being developed.

R. F. Mathis, B. M. Kirsch

INVESTIGATION OF CRYPTOGRAPHIC TECHNIQUES FOR COMPUTER SECURITY AND PRIVACY

The objectives of the research are to develop methods by which sensitive data can be made secure on computers in order to protect individual privacy. These include:

1. Development of the necessary mathematical framework within which the analysis of cryptographic techniques for providing computational security can be performed.

2. Investigation of existing known cryptographic techniques.
3. Determination of the applicability of the known techniques to existing problems in computational security.
4. Investigation of new techniques or extension of existing techniques to provide additional computational security.

F. Stahl (Sponsor: The Ohio State University. Graduate School)

INVESTIGATION OF PATTERN RECOGNITION TECHNIQUES AS APPLIED TO PROBLEMS OF COMPUTER SECURITY

Cryptographic techniques have been proposed and implemented as one means of protecting sensitive information in computer systems. Cryptanalysis, the name for both the scientific study of the security level of proposed crypto systems, and, also, the art of breaking crypto systems, has been almost totally ignored with regard to such systems during their design.

An investigation of the techniques of pattern recognition that can be used to aid in cryptanalysis has been undertaken. Clearly, an encrypted message lays bare before the eyes of the "enemy" all the information that is to be concealed, less, of course, the logical key. All the original redundancy and patterns invoked through the use of the language of the original message are present and merely need to be recognized by the clever analyst.

F. Stahl (Sponsor: Air Force Office of Scientific Research 72-2351)

MAPPING FUNCTIONS FOR INDEPENDENT DATA DEFINITION

The schema-subschema concept has evolved to facilitate independent data definition. The subschema concept has not been present in data base management systems because it is not understood how to interface the schema with a subschema.

The purpose of this research is to study the effect of where the mapping functions are placed, what mapping functions are needed and whether subschemas can be derived from other subschemas.

The goal of this research is not to standardize subschemas, but to discover good design principles for the Data Manager and new mapping algorithms for certain classes of data.

H. S. Koch

A MODEL FOR ACCESS AND ACCESS CONTROL IN DATA BASE MANAGEMENT SYSTEMS

The goal of the research is to propose a model in which fundamental issues in the use of data structures, in the development of data access strategies, and in the employment of both data structures

and access strategies for access control purposes may be studied. The model is a generalization of previous models by the investigators as well as the work by others on similar models. It is hoped that by studying the model, data base system problems and issues may be better understood. Furthermore, systems which make use of data models with complicated structure (e.g., network-structured data based on the CODASYL specifications) and systems which use a content-based approach with simple structures (e.g., relational models) may be compared and contrasted in the generalized model.

Mr. Manola, a Visiting Research Associate, is taking a leave-of-absence from the Naval Research Laboratory.

D. K. Hsiao, F. Manola (Sponsor: Naval Research Laboratory)

PAGINATION OF PROGRAMS FOR VIRTUAL MEMORY SYSTEMS

Pagination for virtual memory systems of modular programs with certain restrictions on their control flow properties has been investigated from two points of view: 1) if the module size is greater than the page size, each module is broken up into basic blocks which are then allocated to pages; and 2) if the module size is less than the page size, entire modules are allocated to pages.

Algorithms based on compile-time loop analysis have been developed, for the instruction portion of such programs, which minimize the page fault rate in each of these cases; for a given page size, the effectiveness of the algorithms is a decreasing function of memory size and, for a given memory size, the effectiveness of the algorithms is an increasing function of page size.

A trace-driven simulation using the XPL compiler-compiler shows that application of these algorithms has the expected effect on the page fault rate.

C. R. Foulk, D. H. Y. Su

PARALLELISM IN COMPUTER PROGRAMS

In 1966 Hellerman proposed a computer design with several arithmetic units instructed from a single sequential program to perform several computations in parallel. One of the early results of the present research has been a technique for the minimization of resources required for minimum time evaluation of algebraic expressions on this Hellerman computer. The generalization of this result to branch-free programs depends on a result of Aho, Sethi, and Ullman, that storage conflicts resulting from multiple uses of a variable name can be avoided by systematic renaming of variables. Identification of independent uses

of variable names has been shown feasible in smooth programs; that is, programs composed of basic blocks combined into two-terminal networks by the operations of concatenation, alternation, and iteration. In a smooth program it is possible to find for each computation a set of appropriate basic blocks. Two computations either share such a set or have disjoint sets. The technique for resource minimization of branch-free programs applies to each set of appropriate basic blocks, and thus generalizes to smooth programs. Algorithms for the automatic application of the technique are being devised.

C. R. Foulk, O. C. Juelich

THE PL/X SYSTEMS PROGRAMMING LANGUAGE

The PL/X language is a high-level systems programming language for the DEC-10 computer. Its syntax is, to a large degree, that of PL/I. It also possesses some of the features available in IBM's systems programming language, PL/S (among them, the DO-CASE Construct). Being specifically designed for the DEC configuration, the PL/X language also allows specifications for octal constants (as opposed to binary bit strings) and provides in-line facility for MACRO-10 code (DEC's assembler language). The PL/X compiler runs on the IBM 370 computer. The compiled object programs are physically transported to the DEC-10 system for execution. The run-time monitor PLXMON, which supervises the execution, has recently been rewritten and upgraded. Consequently, the manual entitled The PL/X Compiler - Subsystem Writers' Manual is being replaced with an updated version.

Plans are made to upgrade the PL/X compiler and the language reference manual. The incorporation of some diagnostic facility into the compiler will also be included.

D. K. Hsiao, B. J. Kammerer, W. Young (Sponsor: The Ohio State University. Instructional and Research Computer Center and Department of Computer and Information Science)

PROVING PROGRAM CORRECTNESS ON COMPLEX FILE STRUCTURES

By using the inductive assertion approach to prove program correctness, this research will expand upon the work of Hoare, Bustall, and Good, who proposed the assertion approach originally. By expanding the work of Laventhal to a complex, but well defined file structure proposed by Hsiao and Harary, proof of programs operating on this file structure can be inductively determined.

Primitive concepts and other abstractions will be related to the structure so that inductive assertions can be made that will prove the correctness of the program by which the file structure

is manipulated. In doing so it is believed that the procedure for verification of data base management systems programs may one day be automated. It is also believed that by analyzing the assertion to actual program relationship, concise and structured code can be generated and that this code will be error-free.

D. K. Hsiao, B. Horger (Sponsor: Office of Naval Research N00014-67-A-0232-0022)

A SMOOTHNESS INDEX FOR COMPUTER PROGRAM FLOWGRAPHS

The ultimate objective of this research is to define an index of smoothness for computer flowgraphs and then to determine whether or not there is any correlation between this index and other properties of such flowgraphs.

A graph grammar, SG, is defined. The productions of SG correspond to the basic and familiar operations of:

concatenation	=	normal sequencing
alternation	=	if-then-else
iteration	=	looping

The "language" produced by this grammar is the set of all flowgraphs which can be generated by these three operations only. Such a flowgraph is called smooth.

A non-smooth flowgraph contains at least one sub-flowgraph which is not generable by SG. In this case, it is possible to transform the original flowgraph into a smooth one.

Transformation of a non-smooth flowgraph to a smooth one always increases the cardinality of the arcs of the flowgraph. By computing the ratio of the cardinality of the arcs before the transformation to the cardinality of the arcs after the transformation, an index of smoothness S ($0 < S \leq 1$) can be obtained.

A computer program which obtains the index of smothness of any given flowgraph has now been written. This program will be used to obtain the index of smoothness of randomly selected flowgraphs from algorithms published in Communications of the ACM. It will then be determined whether or not there is any correlation between the index of smothness of these flowgraphs and their other properties.

C. R. Foulk, O. Longe

TEACHING DEBUGGING

This project is concerned with the integration of debugging concepts into programming courses. What effect does programming

style or structured programming have on debuggability? What information is necessary to debug a program? What examples best illustrate debugging concepts and procedures? One seminar course was held this year and two more are planned for next year to study this problem. A course outline and set of examples are being developed for beginning programming courses.

R. F. Mathis, C. Hasbrouck

USE OF AN ERROR ANALYSIS CONTROL BLOCK (EACB) TO CONTROL PROMPTING IN A POSTMORTEM DEBUGGING SYSTEM

The most common postmortem debugging aid is the system supplied core dump. To make the information in the dump more useful to the programmer/debugger, a dump analysis prompting system is being developed. In this system the core contents are put on disk storage rather than printed. The programmer/debugger is able to investigate this version of a dump through a time sharing terminal. While placing control blocks and core areas on disk, an error analysis control block (EACB) is created which contains relevant flags and pointers from other control blocks and which will be used to control prompting when the dump is investigated. Prompting messages include as many error analysis and correcting suggestions as possible. As new error causes and error analysis methods are discovered they can be semi-automatically added to the system.

R. F. Mathis, R. J. Atwell, Jr.

X. COMPUTER NETWORKS

DISTRIBUTED OPERATING SYSTEMS FOR COMPUTER NETWORKS

This research is concerned with distributed network operating systems. In the same way microcode is used to define the basic machine behavior, the operating system is used to implement a meta-machine. The function of an operating system is then to unify, enrich and control the hardware. In the context of network, an operating system is used in the same manner. The network operating system is used to minimize the user-visible differences between host systems. A conventional operating system controlling a uni-processor or a localized multi-processor is modular, layered, with no redundancy. This type of structure offers very little possibilities in the construction of a network operating system. Standard systems have been designed on the assumption that every module is directly accessible without delay. However the concept of distributed operating system implies that modules may be geographically far apart and thus not immediately available to the controller. In this study a new approach is considered for the design of operating systems. Also, in a conventional operating system, there is no redundancy whereas, in the proposed system, redundancy is a key factor for permanency. This feature implies that a major decision of the concept of operating system has to be undertaken to enable multiple copies of the same module to interact in order to achieve a common goal.

M. T. Liu, S. Fournier

EMULATION OF COMPUTER NETWORKS BY MICROPROGRAMMABLE MICROCOMPUTERS

The advent of low-cost, sophisticated, microprogrammable, LSI microprocessors has renewed interest in multiple-computer systems. This research suggests a method of implementing microprammable microcomputer systems as a sophisticated tool (emulators) for decreasing the economical risk involved in development of a large computer network. Two levels of emulation are proposed for different networking configurations. The Intel 8080 microprocessors are selected for implementing the systems described. As a special case where mutually exclusive data are processed, the proposed systems can be changed to parallel processing systems for parallel execution of computer programs at different levels.

M. T. Liu D. Cohen (Sponsor: National Science Foundation. GN 534.1 and GN 27458)

LOCAL COMPUTER NETWORK

The objective of this project is to implement a local computer network interconnecting IBM 370/165, DECsystem-10 and Micro 1600/21 computer systems. The interconnected systems can serve as a basis for the study of computer network and facilitate better utilization of all the computers involved.

A configuration with low cost and adequate performance has been proposed. This configuration calls for an utilization of the Micro 1600 as a communication processor, a synchronous communication interface between the Micro 1600 and IBM 370/165 and an asynchronous communication interface between the Micro 1600 and DEC System-10. The particular interconnection not only will impact minimally the existing operating systems of the computers involved but also will rely exclusively off-the-shelves hardware components. A cost estimation of the needed hardware components and software development is also included in the proposal.

A design of the software network system is underway.

D. K. Hsiao, B. Farah, T. Rodeheffer (Sponsor: The Ohio State University. Instructional and Research Computer Center and Department of Computer and Information Science)

MIXED DATA/VOICE TRANSMISSION SYSTEM FOR COMPUTER COMMUNICATIONS

This research is concerned with the development of a feasible and inexpensive method for computer communications by embedding digital data signals into a collective set of analog voice channels during the inactive periods of voice conversation. There are two ways of doing this:

1. interleaving or embedding the digital signal into the voice circuit during the silent period of the voice conversation; or
2. transmitting concurrently or superimposing the digital signal onto the voice signal.

Minicomputer/microcomputer technology is used in the control of the digital signal embedding procedure. The goals of the research are:

1. to lower the link cost for computer communications;
2. to improve the usage efficiency of bandwidth in the existing voice communication network.

M. T. Liu, J. T. Wang

OPTIMAL PROCESS ALLOCATION IN DISTRIBUTED HETEROGENEOUS COMPUTER NETWORKS

This research is concerned with optimal process allocation in distributed heterogeneous computer networks so as to minimize hardware/software redundancy. A mathematical model is proposed for analysis using integer and dynamic programming which allows different service rates for each process. A simulation model is to be implemented to verify the results. It is expected that the goal of minimizing redundancy can not only increase overall system efficiency, but can also self-tune the system according to the constantly changing request pattern.

M. T. Liu, T. T. Cheng

VARIABLE-LENGTH MESSAGE TRANSMISSION FOR DISTRIBUTED LOOP COMPUTER NETWORKS

This research is concerned with the development of new protocols for variable-length message transmission in distributed loop computer networks. An improved variable-length message transmission scheme is proposed, and a conceptual model is developed of its operation. The model illustrates a method by which the nodal interface can delay incoming messages by hardware buffering just long enough for an outgoing message to be placed on the loop; advantage is taken of gaps between messages to clear out the delay buffer and to make room for future outgoing messages. Also, some possible hardware realizations of the model are developed. The new scheme has the effects of:

1. reducing the time a message must wait queued at the node before being transmitted,
2. increasing the utilization of the loop, and
3. reducing variations in the instantaneous load factor of the loop.

It is expected that the proposed scheme will yield a transmission technique superior to fixed-length framing currently used by most loop computer networks.

M. T. Liu, C. C. Reames (Sponsor: National Science Foundation. GN 534.1 and GN 27458)

XI. JOINT PROGRAMS

INVESTIGATION OF COMPUTER SCIENCE EDUCATION IN LATIN AMERICA

An investigation has been initiated into the present level of Computer Science Education in Latin America. Of particular interest is the scope of the programs at Latin American Universities and the extent to which these universities are meeting the demand for computer specialists as a result of the rapid increase in the number of computers being utilized in their respective countries. Data will be collected through use of questionnaires and a field trip. Analysis will be performed upon this information and recommendations will be made for future development. The results of this study will be published.

Questions asked are concerned with the present level of programs in the computer area (including numbers of students, level of studies, placement of graduates, language requirements, hardware, and courses offered), level of faculty (including level of education, country of education, and research interest), future plans for expansion of programs, and sources of support (both financial and technical).

A field trip to South America will be undertaken during the latter part of this year. In addition to gathering information about the educational programs the principle investigator will advise and assist in the development of graduate studies and research in the Computer Science Department of Universidad Simon Bolivar, Caracas, Venezuela.

F. Stahl, A. Pumarada (Sponsored in part: UNESCO. 610.247)

APPENDIX A

COMPUTER AND INFORMATION SCIENCE COURSE LISTING

BY NUMBER AND TITLE

100 Computers in Society

201 Elementary Digital Computer Programming

211 Computer Data Processing I

212 Computer Data Processing II

221 Programming and Algorithms I

222 Programming and Algorithms II

294 Group Studies

311 Introduction to File Design and Analysis

411 Design of On-Line Systems

422 Topics in Computing for Engineers

494 Group Studies

505 Fundamental Concepts of Computer and Information Science

509 Survey of Computer and Information Science for High School Teachers

541 Survey of Numerical Methods

542 Introduction to Computing in the Humanities

543 Intermediate Digital Computer Programming

548 Digital Computer Programming for High School Teachers

549 Numerical Analysis for High School Teachers

550 Introduction to Information Storage and Retrieval

555 Survey of Programming Languages

594 Group Studies

610 Principles of Man-Machine Interaction

640 Numerical Analysis

641 Computer Systems Programming I

642 Numerical Linear Algebra

643 Linear Optimization Techniques in Information Processing

644 Advanced Computer Programming

652 Modeling of Information Systems

675 Digital Computer Organization

676 Minicomputer Evaluation and Selection

677 Computer Networks

680 Data Structures

693 Individual Studies

694 Group Studies

705 Mathematical Foundations of Computer and Information Science

706 Information Theory in Behavioral Science

712 Man-Machine Interface

720 Introduction to Linguistic Analysis

726 Theory of Finite Automata

727 Turing Machines and Computability

728 Topics in Theory of Computing

730 Basic Concepts in Artificial Intelligence

735 Statistical Methods in Pattern Recognition

740 Computer Systems Programming II

741 Comparative Operating Systems

745 Numerical Solution of Ordinary Differential Equations

746 Advanced Numerical Analysis

750 Modern Methods of Information Storage and Retrieval

751 Fundamentals of Document-Handling Information Systems

752 Techniques for Simulation of Information Systems

753 Theory of Indexing

754 Language Processing for Information Storage and Retrieval

755 Programming Languages

756 Compiler Design and Implementation

760 Selected Topics in the Mathematics of Information Handling

765 Theory of Management Information Systems

775 Advanced Computer Organization

780 File Structures

781 Aspects of Computer Graphics Systems

788 Intermediate Studies in Computer and Information Science

793 Individual Studies

794 Group Studies

797 Interdepartmental Seminar

805 Information Theory in Physical Science

806 Cellular Automata and Models of Complex Systems

812 Computer and Information Science Research Methods

820 Computational Linguistics

835 Special Topics in Pattern Recognition

840 Operating System Implementation

845 Numerical Solution of Partial Differential Equations

850 Theory of Information Retrieval I

851 Theory of Information Retrieval II

852 Design and Analysis of Information Systems Simulation

855 Formal Theory of Programming Languages

865 Seminar on Socio-Psychological Aspects of the Information Sciences

880 Advanced Theory of Computability

888 Advanced Studies in Computer and Information Science

889 Advanced Seminar in Computer and Information Science

899 Interdepartmental Seminar

994 Group Studies

999 Research

APPENDIX B

COMPUTER AND INFORMATION SCIENCE FACULTY

Marshall C. Yovits, Ph.D., (Yale University).
Professor and Chairman of Department of Computer and Information Science and Professor of Electrical Engineering. Director, C.I.S. Research Center. Information systems, theory of the flow of information, self-organizing systems.

Ranko Bojanic, Ph.D., (Mathematical Institute of the Serbian Academy of Science).
Professor of Computer and Information Science and Professor of Mathematics. Mathematical analysis, theory of approximation.

Richard I. Hang, M.S., (The Ohio State University).
Professor of Computer and Information Science and Professor of Engineering Graphics. Computer graphics, engineering application of computers.

Clyde H. Kearns, M.S., (The Ohio State University).
Professor of Computer and Information Science and Professor of Engineering Graphics. Computer graphics, engineering application of computers.

Robert B. McGhee, Ph.D., (University of Southern California).
Professor of Computer and Information Science and Professor of Electrical Engineering. Control theory, switching theory, logical design.

Harold B. Pepinsky, Ph.D., (University of Minnesota).
Professor of Computer and Information Science and Professor of Psychology. Clinical and socio-cultural psychology.

Roy F. Reeves, Ph.D., (Iowa State University).
Professor of Computer and Information Science and Professor of Mathematics. Director, Instruction and Research Computer Center. Numerical analysis and programming.

Jerome Rothstein, A.M., (Columbia University).
Professor of Computer and Information Science and Professor of Biophysics. Informational problems in science, methodology, biocybernetics.

Charles Saltzer, Ph.D., (Brown Univeristy).
Professor of Computer and Information Science and Professor of Mathematics. Coding theory, numerical analysis, automata theory.

Hugh C. Atkinson, M.A., (University of Chicago).
Associate Professor of Library Administration and Director of Libraries. Library on-line automation, data processing for libraries, technical service processing models.

Kenneth J. Breeding, Ph.D., (University of Illinois).
Associate Professor of Computer and Information Science and Associate Professor of Electrical Engineering. Computer organization and Switching Theory.

Balakrishnan Chandrasekaran, Ph.D., (University of Pennsylvania).
Associate Professor of Computer and Information Science. Pattern recognition and artificial intelligence, learning automata theory, finite memory decision theory, and game theory.

Ronald L. Ernst, Ph.D., (University of Wisconsin).
Associate Professor of Computer and Information Science and Associate Professor of Psychology. Human performance theory and engineering, complex information processing and systems evaluation.

Clinton R. Foulk, Ph.D., (University of Illinois).
Associate Professor of Computer and Information Science. Programming languages, systems programming, programming heuristics.

David K. Hsiao, Ph.D., (University of Pennsylvania).
Associate Professor of Computer and Information Science. Systems programming, information storage and retrieval systems, file systems, data base management systems, access control and privacy protection of data, data definition language and processor, system architectures.

Douglas S. Kerr, Ph.D., (Purdue University).
Associate Professor of Computer and Information Science. Numerical analysis and programming.

Gerald J. Lazorick, Ph.D., (State University of New York at Buffalo).
Associate Professor of Computer and Information Science and Associate Professor of Library Administration. Director, Mechanized Information Center. Information storage and retrieval, library systems: design and analysis.

Ming-Tsan Liu, Ph.D., (University of Pennsylvania).
Associate Professor of Computer and Information Science. Computer organization, switching and automata theory, mathematical programming, computer architecture, pseudo-Boolean programming, threshold logic.

Anthony E. Petrarca, Ph.D., (University of New Hampshire).
Associate Professor of Computer and Information Science. Automatic indexing, chemical structural information processing, automated search systems, other aspects of information storage and retrieval.

James B. Randels, Ph.D., (The Ohio State University).
Associate Professor of Computer and Information Science and Assistant Director, Learning Resources Computer Center. Computer operating systems and utilities, telecommunications applications, subroutine libraries, programming languages.

James E. Rush, Ph.D., (University of Missouri).
Adjunct Associate Professor of Computer and Information Science. Indexing theory, automated language processing, organization of information, and parallel processing.

Celianna I. Taylor, B.S.L.S., (Graduate School of Library Science, Case-Western Reserve University).
Senior Research Associate and Associate Professor of Library Administration. Information dissemination systems, information centers, library systems and management.

Lee J. White, Ph.D., (University of Michigan).
Associate Professor of Computer and Information Science and Associate Professor of Electrical Engineering. Mathematical programming, data structures, organization of information.

Ronald L. Wigington, Ph.D., (University of Kansas).
Adjunct Associate Professor of Computer and Information Science and Director of R. & D., Chemical Abstracts Service. Computer system design.

Alan W. Biermann, Ph.D., (University of California, Berkeley).
Assistant Professor of Computer and Information Science. Theory of computer systems, formal systems.

H. William Buttelmann, Ph.D., (University of North Carolina).
Assistant Professor of Computer and Information Science. Automata theory, computer architecture and programming languages.

Thomas G. DeLutis, Ph.D., (Purdue University).
Assistant Professor of Computer and Information Science. Design and evaluation of information systems, systems programming.

Harvey S. Koch, Ph.D., (Pennsylvania State University).
Assistant Professor of Computer and Information Science. Data definition language, data base management, programming languages and compiler design.

Frederick S. Koehl, Ph.D., (The Ohio State University).
Adjunct Assistant Professor of Computer and Information Science and Math Analyst, Instruction and Research Computer Center. Sorting techniques, topological groups, compiler design.

Anthony P. Lucido, Ph.D., (Iowa State University). Appointed Autumn 1974.
Assistant Professor of Computer and Information Science. Computer architecture, compiler design, interactive computer graphics.

Robert F. Mathis, Ph.D., (The Ohio State University).
Assistant Professor of Computer and Information Science and Assistant Dean and Secretary of the Graduate School. Programming languages, numerical analysis.

Daniel J. Moore, Ph.D., (University of Kansas). Appointed Autumn 1974.
Assistant Professor of Computer and Information Science. Complexity theory, recursion theory, artificial intelligence.

Norman K. Sondheimer, Ph.D., (University of Wisconsin). Appointed Autumn 1974.
Assistant Professor of Computer and Information Science. Natural language processing within artificial intelligence, information storage and retrieval, systems programming.

Frederick A. Stahl, Ph.D., (University of Illinois).
Assistant Professor of Computer and Information Science. Computational security, cryptography, information retrieval, computers in the humanities, and in the law, artificial intelligence.

Stuart H. Zweben, Ph.D., (Purdue University). Appointed Autumn 1974.
Assistant Professor of Computer and Information Science. Programming languages, compilers, data structures, operating systems.

Donald L. Kalmey, M.S., (The Ohio State University).
Counselor.

Ernest Staveley, B.S., (U.S. Naval Postgraduate School).
Administrative Assistant and Assistant Director, C.I.S. Research Center.

APPENDIX C

COMPUTER AND INFORMATION SCIENCE SEMINAR SERIES

August 9, 1973 "The Use of Probabilistic Automata as Models of Some Aspects of Human Memory," Stephen Hepler, Ph.D. Candidate, The Ohio State University.

August 23, 1973 "Human Learning of Cyclic Finite State Languages," James Beug, Ph. D. Candidate, The Ohio State University.

October 4, 1973 "Computer Enhancement of Electron Micrographs," Burn L. Lewis, University of California, Berkeley.

October 18, 1973 "Computer Facilities at Ohio State," Weldon E. Ihrig, Chairman, Computer Center Coordinating Committee; Roy F. Reeves, Director, Instruction and Research Computer Center; Robert F. Mathis, Chairman, Computer and Information Science Computer Committee, The Ohio State University.

October 25, 1973 "Programming at IBM - The Programming Process from Product Design Through Product Support," Larry Foster, Programming Center Manager, IBM Systems Development Division, Gaithersburg, Maryland.

November 1, 1973 "Computers and Medieval Data Processing," Thomas G. Whitney, Instructional and Research Computer Center, The Ohio State University.

November 6, 1973 "Up in the Air about Data Base Management?" Captain Thomas Carroll, Chief of Information Systems Section, Wright-Patterson Air Force Base.

November 8, 1973 "Computer Technology and the Problem of Privacy," Philip M. Burgess, HEW Committee on Privacy and Automated Personal Data Systems, Mershon Professor of Policy Science, Professor of Political Science, The Ohio State University.

November 15, 1973 "Time Flow Mechanisms for Digital Logic Simulation," Cliff W. Hemming, Jr., Bell Telephone Laboratories, Columbus, Ohio.

November 20, 1973 "Formal Studies in Music Composition," Otto E. Laske, Studio for Electronic Music, University of Utrecht, Utrecht, Netherlands.

November 29, 1973 "Programming and Debugging: There Must be a Better Way," Alan W. Biermann, Assistant Professor, The Ohio State University.

January 10, 1974 "Problems in Interactive Bibliographic Files,"

Frederick G. Kilgour, Executive Director, The Ohio College Library Center, Columbus, Ohio.

January 17, 1974 "Software Physics," Maurice H. Halstead, Professor of Computer Sciences, Purdue University.

January 24, 1974 "Symmetric Ternary Radix for Computer Arithmetic," Paul S. Y. Chan, Ph.D. Graduate, The Ohio State University.

January 31, 1974 "The Interactive Generation of Facial Images on a CRT Using a Heuristic Strategy," Mark L. Gillenson, Ph.D. Candidate, Department of Computer and Information Science; and Graduate Research Associate, Computer Graphics Research Group, The Ohio State University.

February 7, 1974 "Development of a Time-Sharing Utility," Dr. John R. Goltz, Senior Vice President, Compu-Serv Network, Columbus, Ohio.

February 19, 1974 "The Application of Some Graph Theoretical Methods of Semantics Information Processing," Stephen L. Taylor, Computer Science Department, Northwestern University.

February 21, 1974 "The Internal Structure of Algorithms," Stuart H. Zweben, Ph.D. Candidate, Computer Sciences Department, Purdue University.

February 26, 1974 "Extraction of Speech Formants Using a Micro-programmed Parallel-Processor System," Michael J. Knudsen, Department of Computer Science, Carneigie-Mellon University.

February 27, 1974 "The Computational Complexity of Oracular Enumeration Procedures," Daniel J. Moore, Computer Science Department, University of Kansas.

February 28, 1974 "Some Problems of Spatial Reference for Language Understanding Systems," Norman Sondheimer, Ph.D. Candidate, Computer Sciences Department, University of Wisconsin.

March 5, 1974 "Design and Implementation of a Data Conversion Language for Networks," Michael Schneider, Computer Science Department, University of Wisconson.

March 7, 1974 "Load Sharing of Batch Jobs: A Queueing Model of a Computer Network," Carl Landwehr, Computing Center, University of Michigan.

March 12, 1974 "A Model for the Evaluation and Optimization of File Designs," Bing Yao, Computer and Communication Sciences Department, University of Michigan.

March 12, 1974 "A Model for the Evaluation and Optimization of File Designs," Bing Yao, Computer and Communication Sciences Department, University of Michigan.

March 14, 1974 "The Grammatical Inference of Formal Languages," Jan Beaujon, North Carolina State University at Raleigh.

March 26, 1974 "A Homogeneous Network for Data Sharing," R. W. Peebles, Computer Communications Network Group, University of Waterloo.

March 27, 1974 "On the META-T Compiler-Compiler," Anthony Lucido, Iowa State University.

March 28, 1974 "A Graphics Extension to SNOBOL4 for Recognition and Manipulation of Line Drawings," Linda Shapiro, University of Iowa.

April 2, 1974 "Computer Display of Curved Surfaces," Edwin Catmull, University of Utah, Salt Lake City, Utah.

April 4, 1974 "Computing Reliability of Complex Systems," Aron Rosenthal, Department of Electrical Engineering and Computer Science, University of California at Berkeley.

April 8, 1974 "Biographical Information," William E. Riddle, Assistant Professor, Computer and Communication Sciences Department, University of Michigan.

April 22, 1974 "The Design of an Integrated Medical Record Information System," Kevin O'Kane, Assistant Professor of Computer Science and Assistant Professor of Obstetrics and Gynecology, Pennsylvania State University.

April 25, 1974 "Some Results on Computability of Phrase Structure Languages," H. William Buttelmann, Assistant Professor of Computer and Information Science, The Ohio State University.

May 2, 1974 "Protecting Ones Proprietary Interest in Software - Which Way to Turn?" James A. Sprowl, Lecturer, School of Law, Northwestern University, Partner, Lawfirm of Mason, Kolehmainen, Rathburn and Wyss.

May 23, 1974 "Fourth Generation Computer Architecture," Harold S. Schwenk, Honeywell Information Systems.

APPENDIX D

RELATED ACTIVITIES OF THE STAFF OF

COMPUTER AND INFORMATION SCIENCE RESEARCH CENTER

J. G. Abilock presented "Comparison of Relational Data Management Systems" (Co-author: V.K.M. Whitney), and "A Quantitative Measure for the Information Contained in a Data Set or a Message" (Co-author: M.C. Yovits) at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

J. A. Aitken presented the paper "Performance Evaluation of Information Systems: Data Base Characterization Facilities" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

H. C. Atkinson was a consultant for Library Automation at Penn State University, College Park, Pennsylvania, July 19-21, 1973.

H. C. Atkinson presented "Will Computers Depersonalize Library Service?" State University of New York, Albany, September 21, 1973.

H. C. Atkinson served as Panel Chairman of "Data Banks and Libraries" section of EDUCOM Fall 1973 Council Meeting and Conference on "Facts and Futures: What's Happening Now in Computing for Higher Education" at the Educational Testing Service, Princeton, New Jersey, October 9-11, 1973.

H. C. Atkinson conducted a seminar for library staff at the University of Tennessee, Knoxville, Tennessee, May 15, 1974.

R. J. Atwell, Jr. presented the paper "Use of an Error Analysis Control Block (EACB) to Control Prompting in a Post Mortem Debugging System" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

A. W. Biermann; F. Petry attended the Third International Joint Conference on Artificial Intelligence, Stanford University, Stanford, California, August 20-23, 1973.

A. W. Biermann; R. Baum; R. Krishnaswamy; F. Petry made a 10 minute color movie entitled Autoprogrammer, March 1974.

A. W. Biermann gave seminar talks entitled "Constructing Programs from Example Computations" and showed the movie Autoprogrammer at Duke University and the University of Texas.

H. W. Buttelmann presented "Servantic Directed Translation of Context Free Languages" at the annual meeting of the Association for Computational Linguistics, at the University of Michigan, August 2, 1973.

H. W. Buttelmann presented a paper entitled "Applications in Automated Language Processing" at the 1974 Annual Meeting of the Ohio Academy of Sciences at Wooster, Ohio, April 24-26, 1974.

J. Chandler presented the paper "Performance Evaluation of Information Systems: Simulation Time Facilities and Statistics" at the Computer Science Conference, Detroit, Michigan, February 11,-14, 1974.

B. Chandrasekaran was a member of the panel on "Pattern Recognition: Bridging the Gap Between Theory and Implementation" as part of the International Joint Conference on Pattern Recognition, Washington, D. C., October 30, 1973.

B. Chandrasekaran was the organizer and chairman of the Norbert Wiener Commemorative Symposium, Boston, Massachusetts, November 7, 1973.

B. Chandrasekaran was Chairman of the session on Artificial Intelligence I, at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

B. Chandrasekaran presented "A Deterministic Finite Memory Algorithm for the Symmetric Hypothesis Testing Problem" at the IEEE International Information Theory Symposium, Ashkelon, Israel, June 28, 1974.
(Co-author: C.C. Lam)

B. Chandrasekaran has been elected a Vice-President of the Systems, Man and Cybernetics Society of the Institute of Electrical and Electronics Engineers.

B. Chandrasekaran was a reviewer for the IEEE Transactions Information Theory Information and Control, Computing Reviews, 1973-1974.

B. Chandrasekaran reviewed proposals for the National Science Foundation Office of Computing Activities and NSF Engineering Division for 1973-1974.

T. G. DeLutis presented "Recent Developments in Data Description Languages" for the Central Ohio Chapter Association for Computing Machinery, Columbus, Ohio, October 8, 1973.

T. G. DeLutis presented "Performance Evaluation of Information Systems: The AUXSYM Simulator" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

C. R. Foulk presented "Structured Programming in Unstructured Languages" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

C. R. Foulk served on the organizing committee of the annual meeting of the Association for Computing Machinery, Special Interest Group on Computer Science Education, Detroit, Michigan, February 14-15, 1974.

D. K. Hsiao was invited to attend the IBM Data Security Sympsium and participated in the workshop on issues related to data security, New York, October 29 - November 1, 1973.

D. K. Hsiao presented the paper "A Unified Approach to Structure, Access and Update in Data Base Systems" at the European Computing Congress held at Brunel University, London, England, May 14, 1974.

D. K. Hsiao and nine undergraduate and graduate students from the Department of Computer and Information Science have recently completed the design and implementation of the first PL/1-like compiler language for the DECsystem-10 computer. The compiler language is called the PL/X whose syntax is to a large extent the syntax of the PL/1.

S. James presented the paper "A Sequential Approach to Automatic Document Classification and Retrieval" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.
(Co-authors: L.J. White, J. Rush, D. Westbrook, G. Jones, G. Kar)

D. L. Kalmey presented "Numerical Experimentation with Routines for Solving Non-linear Systems of Equations" at the Computing Machinery Annual Conference, Atlanta, Georgia, August 27-29, 1973.
(Co-authors: D.S. Kerr, L.J. White)

D. L. Kalmey presented the paper "Routines for Solving Nonlinear Systems of Equations" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.
(Co-authors: D.S. Kerr, L.J. White)

C. Kearns has been elected to serve a three-year term as circulation manager-treasurer of the Engineering Design Graphics Jornal.

D. S. Kerr attended the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

D. S. Kerr served as Chairman of the annual meeting of the Association for Computing Machinery, Special Interest Group on Computer Science Education in Detroit, Michigan, February 14-15, 1974.

D. S. Kerr attended the ACM-SIGFIDET Workshop on Data Description, Access and Control, Ann Arbor, Michigan, May 1-3, 1974.

D. S. Kerr was a panelist for "Research in Data Security - Policy and Projects" at the National Computer Conference, Chicago, Illinois, May 6-10, 1974.

D. S. Kerr was appointed Chairman of the Central Ohio Chapter of the Association for Computing Machinery.

B. M. Kirsch presented the paper "Improved Error Diagnostics for IBM System/370 OS/MVT Assembler Program Dumps" at the Computer Science Conference, Detroit Michigan, February 11-14, 1974.

H. S. Koch presented the paper "Data Base Reorganization Tools" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

H. S. Koch was accepted as a member of the Subschema Task Group of the CODASYL Data Definition Language Committee.

M. J. Lee presented "Design Specifications for an Information Retrieval System Management Game" at the Association for Computing Machinery Annual Conference, Atlanta, Georgia, August 27-29, 1973. (Co-author: G.J. Lazorick)

O. Longe presented the paper "An Index of Smoothness of Program Flowgraphs" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

R. F. Mathis chaired one of the Research Abstracts Sessions at the Association for Computing Machinery Annual Conference, Atlanta, Georgia, August 27-29, 1973.

R. F. Mathis attended the organizational meeting of Local Users Group of DECUSO, Pittsburgh, Pennsylvania, September 11, 1973.

R. F. Mathis was the chairman of the ASEE/OSU Critchfield Effective Teaching Institute, Hocking Hills State Park, Ohio, September 18-20, 1973.

R. F. Mathis was appointed Secretary of the Computer Science Conference Board at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

R. F. Mathis presented the paper "Program Debugging in a General Program Development Model" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

R. F. Mathis presented "Teaching Debugging" at the Fourth SIGCSE Technical Symposium, Detroit, Michigan, February 14-15, 1974.

R. F. Mathis was the leader of the ASEE Compus Activities Coordinators' Workshop, Mansfield, Ohio, March 1-2, 1974.

R. F. Mathis attended the ASEE North Central Section Meeting, Ann Arbor, Michigan, April 19-20, 1974.

R. F. Mathis was the session chairman for "Research in Data Security - Policy and Projects" at the National Computer Conference, Chicago, Illinois, May 6-10, 1974.

R. F. Mathis, R. LaRue presented "Computers in the Engineering Classroom" as part of the ASEE Box-lunch Forum, The Ohio State University, Columbus, Ohio, May 6, 1974.

R. F. Mathis attended the CIC Graduate Deans meeting, Ann Arbor, Michigan, May 12-14, 1974.

R. F. Mathis presented "Using the Computer to Teach Design and Testing" at the ASEE Annual Conference, Rensselaer Polytechnic Institute, Troy, New York, June 17-20, 1974.

R. F. Mathis attended the Computer Science Department Chairmen's meeting, Salt Lake City, Utah, June 30 - July 3, 1974.

R. F. Mathis was nominated to honorary membership in TEXNOKOI, an Honorary Activities Fraternity, of the College of Engineering, The Ohio State University.

A. E. Petrarca presented the paper "Algorithmic Control of Index Term Specificity in Double-KWIC Coordinate Indexing Systems" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

F. Petry presented "Program Synthesis from Memory Snapshots of Sample Computations" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.
(Co-author: A.W. Biermann)

F. Petry gave a seminar talk entitled "Inference of Programs from Memory Traces" and showed the movie Autoprogrammer at Wayne State University, March 1974; and at the University of Southwestern Louisiana, April 8, 1974.

A. Pyster presented the paper "The Role of the Nonterminal in Language Generation" at the Computer Science Conferency, Detroit, Michign, February 11-14, 1974.
Co-authors: H.W. Buttelmann, L. Reeker)

J. Rothstein presented "A Cellular Automaton for Pattern Recognition by Parallel Computation" (Co-author: C. Weiman) and "A New Binary Code for Pattern Recognition by Parallel Computation" at the Sagamore Computer Conference on Parallel Processing, August 22-24, 1973 and at the Association for Computing Machinery Annual Conference, Atlanta, Georgia, August 29, 1973.

J. Rothstein presented the paper "The Rothstein Straight Line Code as a Formal Language" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

J. Rothstein presented "Fission Products as an Energy Resource" before the American Physical Society, Philadelphia, Pennsylvania, March 25-28, 1974.

F. A. Stahl presented "Semantics Modeling for Deductive Question-Answering" at the Third International Joint Conference on Artificial Intelligence, Stanford University, August 1973.

F. A. Stahl attended the NBS conference on Privacy and Security in Computer Systems, November 19-20, 1973.

F. A. Stahl presented a talk entitled "Conservation, Computers, and You" at the Computer Science Conference, Wayne State University, February 1974.

F. A. Stahl was awarded an Ohio State University Research Grant to investigate cryptographic techniques for computer security and privacy, June 1974.

S. M. Strong had her paper "An Algorithm for Generating Structural Surrogates of English Text" chosen as the winner of the 1973 American Society for Information Science (ASIS) Student Member Paper Contest.

C. I. Taylor became a member of the Board of Trustees of the Columbus Regional Information Service (CRIS) serving the Columbus Area Chamber of Commerce and the Columbus area citizens, March 15, 1974.

C. I. Taylor served as a consultant to the Directive Teaching Instructional Materials System (DTIMS) project on the development of a computerized data base, the Department of Exceptional Children, The Ohio State University, May - June 1974.

C. I. Taylor attended the Planning Conference for NCEMMH/ALRC National Information Retrieval System, The Ohio State University, June 6-7, 1974.

C. I. Taylor prepared the work statement comprising the technical content of a Request for Proposal (RFP) regarding an information retrieval system for a national audience, serving the handicapped and requiring access to special education materials, June 1974.

W. M. Wagner presented the paper "Asympototes of a Differential Equation of Ion Recombination" at the Association for Computing Machinery Annual Conference, Atlanta, Georgia, August 27-29, 1974. (Co-author: H.R. Wood)

W. M. Wagner presented the paper "Asymptotes of a Differential Equation of Ion Recombination of an Arbitrary Number of Different Molecules" at the Computer Science Conference, Detroit, Michigan, February 11-14, 1974.

C. F. R. Weiman presented the paper "A Cellular Automaton for Pattern Recognition by Parallel Computation" at the Association for Computing Machinery Annual Conference, Atlanta, Georgia, August 27-29, 1973.

M. C. Yovits organized the Research Abstracts Sessions and chaired one of the sessions at the Association for Computing Machinery Annual Conference, Atlanta, Georgia, August 27-29, 1973.

M. C. Yovits presented a paper on "A Quantitative Measure for the Information Contained in a Data Set or a Message" at the Association for Computing Machinery Annual Conference, Atlanta, Georgia, August 27-29, 1973.
(Co-author: J.G. Abilock)

M. C. Yovits visited Rutgers University as part of the Visiting Committee for Mathematical Sciences, October 15-17, 1973.

M. C. Yovits presented an invited paper entitled "A Theoretical Framework for the Development of Information Science" at the International Federation of Documentation Committee on Research on the Theoretical Basis of Information (FID/RI) meeting, Moscow, U.S.S.R., April 24-26, 1974.

M. C. Yovits spoke on "Information Science Research in the Soviet Union" at a joint meeting of the COCACM and COACM, Battelle Memorial Institute, Columbus, Ohio, May 13, 1974.

M. C. Yovits was a panelist on "Information Transfer Today and Tomorrow" at the American Society for Engineering Education National Conference, Rensselaer Polytechnic Institute, Troy, New York, June 17-20, 1974.

M. C. Yovits accepted an invitation to serve on an Advisory Committee for the Department of Computer and Information Science at the University of Pennsylvania.

APPENDIX E

PUBLICATIONS OF THE COMPUTER AND INFORMATION SCIENCE RESEARCH CENTER STAFF[1]

ABILOCK, J. G. An evaluation of relational data base management systems. Computer Science Department, General Motors Research Laboratories, January, 1974.

ATKINSON, H. C. Changing objectives in research libraries. In: Minutes of the 83rd meeting of the Association of Research Libraries, Chicago, Illinois, January, 1974.

ATKINSON, H. C., compiler Book trade statistics. In: The Bowker Annual of Library and Book Trade Information, 19th ed. New York, R. R. Bowker, 1974.

ATKINSON, H. C. Circulation -- an overview. In: Advances in librarianship, (edited by Melvin J. Voigt) New York, Academic Press, 1974, p. 63-75.

ATKINSON, H. C. The future of the scholarly journal. In: Management problems in serials work, (edited by Peter Spyers-Duran) Greenwood Press, 1974.

ATKINSON, H. C., ed. Prices of U.S. and foreign published materials. In: The Bowker Annual of Library and Book Trade Information, 19th ed. New York, R. R. Bowker, 1974.

CHANDRASEKARAN, B.; REEKER, L. H. Artificial intelligence: A case for agnosticism. IEEE Transactions on Systems, Man and Cybernetics (January, 1974).

CHANDRASEKARAN, B.; JAIN, A. K. Distance functions for independent measurements and finite sample size. In: Proceedings II International Joint Conference on Pattern Recognition, Denmark, August 1974.

CHANDRASEKARAN, B.; JAIN, A. K. Quantization complexity and independent measurements. IEEE Transactions on Computers (January, 1974).

GUDES, E. On the evaluation of Boolean expressions. Journal of Software-Practices and Experience (December, 1973).

---

[1]See Appendix F for publications issued as part of the Computer and Information Science Research Center technical report series.

HSIAO, D. K.; KERR, D. S.; STAHL, F. A. Research on data secure systems. In: Proceedings of the AFIPS Conference, 43, Montvale, New Jersey, 1974, p. 994-996.

HSIAO, D. K. Systems programming -- Preliminary edition. Addison-Wesley Publishing Co., April, 1974.

LI, Y. K. Evaluating information storage and retrieval system - A decision theory approach. Information Storage and Retrieval, 9:5 (May, 1973).

MATHIS, R. E., ed. Research in data security - Policies and projects. In: AFIPS Conference Proceedings, 43, Montvale, New Jersey, AFIPS Press 1974, p. 993-999.

MATHIS, R. F. Teaching debugging. SIGCSE Bulletin, 6:1 (February, 1974) p. 59-63.

ROTHSTEIN, J. Loschmidt's and Zermelo's paradoxes do not exist. Foundations of Physics, 4:1 (January, 1974).

ROTHSTEIN, J. A method for constructing singly even magic squares. The Fibonacci Quarterly, 11:5 (December, 1973).

STAHL, F. On applications of mathematics in cryptography. Notices of the American Mathematical Society, 21:4 (June, 1974).

STAHL, F. On computational security. Coordinated Science Laboratory, University of Illinois, Urbana-Champaign, January, 1974.

STAHL, F.; HSIAO, D. K.; KERR, D. Research on data secure systems. In: Proceedings AFIPS 1974 NCC, 43, Montvale, New Jersey, AFIPS Press, 1974, p. 994-996.

STAHL, F.; CHIEN, R.; WEISSMAN, S. Semantic modeling for deductive question answering. In: Proceedings of the Third International Joint Conference on Artificial Intelligence (August, 1973) p. 345-363.

WHITE, L. J.; CHAKAL, J. S. Bilinear surface representation of radar scattering for aircraft. Electro Science Laboratory, Department of Electrical Engineering, The Ohio State University, July, 1973 (Technical Report 2768-10).

YOVITS, M. C.; WHITTEMORE, B. J. A generalized conceptual development for the analysis and flow of information. Journal of the American Society for Information Science (May-June, 1973).

PAPERS ACCEPTED FOR PUBLICATION

ABILOCK, J. G.; YOVITS, M. C. A semiotic framework for information science leading to the development of a quantitative measure of information. In: Proceedings of the ASIS Annual Conference, Atlanta, Georgia, October, 1974.

ATKINSON, H. C. Extension of new services and the role of technology. Library Trends (October, 1974).

ATKINSON, H. C. Priorities for new librarian of congress. Library Journal (July, 1974).

ATKINSON, H. C. The Ohio State University mechanized information center. In: Proceedings of EDUCOM 1973 Fall Conference.

BIERMANN, A. W.; BAUM, R. I.; PETRY, F. E. Speeding up the synthesis of programs for tracos. IEEE Transactions on Computers.

CHANDRASEKARAN, B. Artificial intelligence: The past decade. Advances in Computers, 13, Academic Press.

CHANDRASEKARAN, B.; JAIN, A. K. Distance functions for independent measurements and finite sample size. II International Joint Conference on Pattern Recognition, Lyrgby, Denmark, 1974.

CHANDRASEKARAN, B.; LAM, C. C. A finite memory deterministic algorithm for the symmetric hypothesis testing problem. IEEE Transactions on Information Theory.

CHANDRASEKARAN, B.; JAIN, A. K. Independence, measurement complexity and classification performance. IEEE Transactions on systems, man and cybernetics.

LIU, M. T.; COHEN, D. Emulation of computer networks by microprogrammable microcomputer. In: Proceedings of the ACM 7th Annual Workshop on Microprogramming.

ROTHSTEIN, J.; BAMMEL, S. E. The number of 9 x 9 Latin squares. Discrete Mathematics.

STAHL, F. Cryptography and computation. In: Encyclopedia of computer science (edited by A. Ralston) Auerback, 1974.

STAHL, F. Cryptography in the age of automation. In: The Proceedings of the IEEE International Sympsium on Information Theory, October, 1974.

WHITE, L. J.; KSIENSKI, A. A. Aircraft identification using a bilinear surface representation of radar data. Pattern Recognition, 6, 1974.

PAPERS SUBMITTED FOR PUBLICATION

BIERMANN, A. W.; KRISHNASWAMY, R. A system for program synthesis from examples. IEEE Systems, Man and Cybernetics 1974 Conference, Dallas, Texas, October, 1974.

CHANDRASEKARAN, B.; GILLENSON, M. L. The generation of human facial images on a CRT using a heuristic strategy. Communications of ACM.

KERR, D. S.; WHITE, L. J.; WANG, P. T. R. Bandwidth reduction of sparse matrices by row and column permutations. SIAM Journal on Numerical Analysis.

KOCH, H. S. A missing component of current data base management systems - Data base reorganization. 1974 ASIS Annual Meeting.

LIU, M. T.; REAMES, C. C. Variable-length message transmission for distributed loop computer networks. Second Annual Symposium on Computer Architecture, Huston, Texas, January, 1975.

WANG, T. R.; KERR, D. S.; WHITE, L. J. Bandwidth reduction of sparse matrices by row and column permutations. SIAM Journal on Numerical Analysis, 1974.

KSIENSKI, A. A.; REPJAR, A. E.; WHITE, L. J. Object identification for multi-frequency radar returns. In: Proceedings of the IEEE, 1974.

APPENDIX F

TECHNICAL REPORT SERIES

1968

YOVITS, M. C.; ERNST, R. L. Generalized information systems: Some consequences for information transfer. October, 1968. 47p. (OSU-CISRC-TR-68-1) (PB-180 929)

FILLMORE, C. J.; LEHISTE, I. Working papers in linguistics no. 2. November, 1968. 128p. (OSU-CISRC-TR-68-3) (PB-182 596)

FRIED, J. B.; LANDRY, B. C.; LISTON, JR., D. M.; PRICE, B. P.; VAN BUSKIRK, R. C.; WASCHSBERGER, D. M. Index simulation feasibility and automatic document classification. October, 1968. 21p. (OSU-CISRC-TR-68-4) (PB-182 597)

ROTHSTEIN, J. Thermodynamics & information: Before, in and beyond quantum mechanics. December, 1968. 21p. (OSU-CISRC-TR-68-5) (PB-183 738)

FINLEY, JR., M. R. The development of a basic language for artificial intelligence. January, 1969. 24p. (OSU-CISRC-TR-68-6) (PB-182 305)

1969

COLOMBO, D. S.; RUSH, J. E. Use of word fragments in computer-based retrieval systems. February, 1969. 7+[9]p. (OSU-CISRC-TR-69-1) (PB-184 104)

WHITE, L. J. Minimum covers of fixed cardinality in weighted graphs. March, 1969. 14p. (OSU-CISRC-TR-69-2) (PB-183 737)

JACKSON, D. M. The construction of retrieval environments and pseudo-classifications based on external relevance. April, 1969. 74p. (OSU-CISRC-TR-69-3) (PB-184 462)

ELLIOT, D. E.; HUANG, S.; LANGENDOEN, D. T.; LEE, P. G.; LEHISTE, I. Working papers in linguistics no. 3. June, 1969. 181p. (OSU-CISRC-TR-69-4)

BRIGGS, G. E. Reaction time and uncertainty in human information processing. March, 1969. 36p. (OSU-CISRC-TR-69-5) (PB-184 135)

WHITE, L. J.; RUSH, J. E. Linear lists for spiro graphs. June, 1969. 69p. (OSU-CISRC-TR-69-6) (PB-194 402)

PETRARCA, A. E.; LAY, W. M. The double KWIC coordinate index. A new approach for preparation of high-quality printed indexes by automatic indexing techniques. April, 1969. 12+[17]p. (OSU-CISRC-TR-69-7)

YOVITS, M. C. Information science: Toward the development of a true scientific discipline. June, 1969. 27p. (OSU-CISRC-TR-69-8) (PB-187 983)

PETRARCA, A. E.; LAY, W. M. The double KWIC coordinate index. II. Use of an automatically generated authority list to eliminate scattering caused by some singular and plural main index terms. August, 1969. 13p. (OSU-CISRC-TR-69-9)

MCCULLOUGH, J. L. The acquisition of information across cultures: I. Persuasive role play, counterargument and attitude change. August, 1969. 18p. (OSU-CISRC-TR-69-10) (PB-197 568)

ERNST, R. L.; YOVITS, M. C. Information science as an aid to decision making. September, 1969. 22p. (OSU-CISRC-TR-69-13) (PB-189 666)

LANDRY, B. C. An indexing and re-indexing simulation model. June, 1969. 50p. (OSU-CISRC-TR-69-14) (PB-198 115)

SALVADOR, R. Automatic abstracting and indexing. June, 1969. 93p. (OSU-CISRC-TR-69-15)

STEVENS, D. W. A computer program for the reduction of flow tables. June, 1969. 97p. (OSU-CISRC-TR-69-16) (PB-189 679)

COLOMBO, D. S. Automatic retrieval systems and associated retrieval languages. 1969. 69p. (OSU-CISRC-TR-69-17) (PB-198 116)

SCHLESSINGER, J. D.; WHITE, L. J. Optimum prefix encoding. August, 1969. 85p. (OSU-CISRC-TR-69-18) (PB-198 117)

DAY, R.; WHITE, L. J. Hebbian neural simulation: Computer program documentation. (OSU-CISRC-TR-69-19) (PB-204 003)

REEKER, L. H. Extended finite state representation of infinite machines. September, 1969. 36p. (OSU-CISRC-TR-69-20) (PB-187 949)

WILLIAMS, N. T.; ERNST, R. L. A computer simulation of human short-term memory. 1969. 62p. (OSU-CISRC-TR-69-22) (PB-197 874)

BEZDEK, R. R. The acquisition of information across cultures: II. Social science research in a different culture. III. Cross-cohort activity and attitude change, by J. L. McCullough. January, 1970. 36p. (OSU-CISRC-TR-69-23) (PB-197 876)

1970

UNKLESBAY, M. K. A one step version of Younger's algorithm for bounded context grammars. 1970. 41p. (OSU-CISRC-TR-70-1) (PB-197 603)

LI, Y. Information structure and optimal policy. September, 1970. 18p. (OSU-CISRC-TR-70-2) (PB-197 605)

DILLON, S. R. Some procedures for finding substitution property partitions, substitution property covers, and cover pairs for finite state sequential machines. 1970. 79p. (OSU-CISRC-TR-70-3) (PB-197 643)

MATHIS, B. A.; WHITE, L. J.; JACKSON, D. M. Stability analysis of term similarities for information classification theory. July, 1970. 79p. (OSU-CISRC-TR-70-4) (PB-195 376)

MCGHEE, R. B.; DILLON, S. R. A paull-unger procedure for substitution property partitions. April, 1970. 16p. (OSU-CISRC-TR-70-5) (PB-192 120)

DAY, R. G.; WHITE, L. J. Study of a random search method for function minimization. March, 1970. 77p. (OSU-CISRC-70-6) (PB-194 404)

PHARES, R.; WHITE, L. J. Identification of circuits in chemical structures. June, 1970. 73p. (OSU-CISRC-TR-70-7) (PB-194 396)

HARALSON, K. M.; WHITE, L. J. Optimal prefix codes for ensembles of N equiprobable messages using a binary alphabet. May, 1970. 105p. (OSU-CISRC-TR-70-8) (PB-197 642)

PETRARCA, A. E.; LAITINEN, S. V.; LAY, W. M. Use of the double KWIC coordinate indexing technique for chemical line notations. 1970. 14+[17]p. (OSU-CISRC-TR-70-9) (PB-198 269)

LAY, W. M.; PETRARCA, A. E. Modified double KWIC coordinate index. Refinements in main term and subordinate term selection. 1970. 11+[11]p. (OSU-CISRC-TR-70-10) (PB-197 567)

LYONS, J. J. The speed-accuracy trade-off in processing different classes of material. 1970. 38p. (OSU-CISRC-TR-70-11) (PB-198 114)

FILLMORE, C. J.; LEHISTE, I.; MELTZER, D.; TATHAM, M. A.; THOMPSON, S. A. Working papers in linguistics no. 6. September, 1970. 132p. (OSU-CISRC-TR-70-12) (PB-194 829)

ROTHSTEIN, J. Information generalization of entropy in physics. February, 1970. 22p. (OSU-CISRC-TR-70-24) (PB-192 128)

JACKSON, D. M. Basis for an improvability measure for retrieval performance. February, 1970. 31p. (OSU-CISRC-TR-70-25) (PB-197 812)

DRACHMAN, B.; EDWARDS, M. L.; FILLMORE, C. J.; LEE, G.; LEE, P.; LEHISTE, I.; ZWICKY, A. M. Working papers in linguistics no. 4 May, 1970. 164p. (OSU-CISRC-TR-70-26) (PB-192 163)

1971

ROTHSTEIN, J. Patterns and algorithms. January, 1971. 8p. (OSU-CISRC-TR-71-1) (PB-197 604)

GROSU, A.; LEE, G. Working papers in linguistics no. 7. February, 1971. [243]p. (OSU-CISRC-TR-71-2) (PB-198 278)

CHANDRASEKARAN, B.; KANAL, L. On linguistic, statistical, and mixed models for pattern recognition. March, 1971. 33,A5p. (OSU-CISRC-TR-71-3) (PB-198 279)

WHITTEMORE, B. An example of the application of generalized information systems concepts to the quantification of information in a decision system: The examination of quantified information flow in an industrial control problem. May, 1971. 51p. (OSU-CISRC-TR-71-4) (PB-202 621)

OSTROM, T. M.; STEELE, C. M.; SMILANSKY, J. Information and attitudes: The effects of information context and perceived discrepancy on attitudes. May, 1971. 23p. (OSU-CISRC-TR-71-5) (PB-202 622)

LI, Y. Equipment replacement models: A generalization and extension. May, 1971. 17p. (OSU-CISRC-TR-71-6) (PB-200 548)

ELLIOTT, D.; GEIS, M.; GROSS, A.; NOBEL, B.; ZWICKY, ANN; ZWICKY, ARNOLD. Working papers in linguistics no. 8. June, 1971. 197p. (OSU-CISRC-TR-71-7) (PB-202 724)

BOND, Z. S.; GREGORSKI, R.; KEREK, A.; LEHISTE, I.; SHOCKEY, L.; WENDELL, M. V. Working papers in linguistics no. 9. July, 1971. 232p. (OSU-CISRC-TR-71-8) (PB-204 002)

JOHNSEN III, A. M. Performance in memory scan task under conditions of fixed versus varied memory sets. November, 1971. 49p. (OSU-CISRC-TR-71-9)

MEHAFFEY III, L. The spectral sensitivity of the turtle Pseudemys scripta elegans. November, 1971. 89p. (OSU-CISRC-TR-71-10)

WILLIAMS, J. D. Stimulus encoding and human information processing. November, 1971. 65p. (OSU-CISRC-TR-71-11)

SANDERS, S. A. A modification of a method of generating random numbers using a combination of two congruential generators. December, 1971. 76p. (OSU-CISRC-TR-71-12)

LANDRY, B. C. A theory of indexing: Indexing theory as a model for information storage and retrieval. December, 1971. 271p. (OSU-CISRC-TR-71-13) (PB-205 829)

OSTROM, T. M.; EDWARDS, J. D.; ROSENBLOOD, L. K. Integration of discrepant information in interpersonal attitudes. December, 1971. 43p. (OSU-CISRC-TR-71-14)

1972

WHITE, L. J.; GILLENSON, M. L. Optimum center location. January, 1972. 69,A6,B32p. (OSU-CISRC-TR-72-1)

MELTZER, D. Speech synthesis by haar functions with comparison to a terminal analog device. January, 1972. 135p. (OSU-CISRC-TR-72-2)

CHANDRASEKARAN, B.; JAIN, A. K. Quantization of independent measurements and recognition performance. March, 1972. 14p. (OSU-CISRC-TR-72-3)

CAMERON, J. S. Automatic document pseudoclassification and retrieval word frequency techniques. March, 1972. 165p. (OSU-CISRC-TR-72-4)

OSTROM, T. M.; SLOAN, L. R.: MC CULLOUGH, J. L. Information and attitudes: The effects of repetition and amount of information. April, 1972. 38p. (OSU-CISRC-TR-72-5) (PB-209 802)

LEHISTE, I.; MELTZER, D.; SHOCKEY, L.; GREGORSKI, R. Working papers in linguistics no. 12. June, 1972. 88p. (OSU-CISRC-TR-72-6) (PB-210 781)

EKONG, V. J. U. Rate of convergence of hermite interpolation based on the roots of certain Jocabi polynomials. June, 1972. 54p. (OSU-CISRC-TR-72-7) (PB-211 237)

WEIMAN, C. F. R.; ROTHSTEIN, J. Pattern recognition by retina-like devices. July, 1972. 154p. (OSU-CISRC-TR-72-8) (PB-211 238) (PB-214 665)

CHANDRASEKARAN, B.; REEKER, L. H. Artificial intelligence - a case for agnosticism. August, 1972. 28p. (OSU-CISRC-TR-72-9)

ROTHSTEIN, J. Loschmidts's and Zermelo's paradoxes do not exist. October, 1972. 9p. (OSU-CISRC-TR-72-10) (PB-213 712) (PB-214 666)

YOVTIS, M. C.; CHANDRASEKARAN, B. Artificial intelligence. December, 1972. 39p. (OSU-CISRC-TR-72-11)

HSIAO, D. K.; DENICOFF, M.; BERGART, J. G. An annotated and cross-referenced bibliography on computer security and access control in computer systems. November, 1972. 57p. (OSU-CISRC-TR-72-12)

FOULK, C. R.; JUELICH, O. C. Smooth programs and languages. November, 1972. 16p. (OSU-CISRC-TR-72-13) (PB-218 805)

YOVITS, M. C.; WHITTEMORE, B. A generalized conceptual development for the analysis and flow of information. December, 1972. 31p. (OSU-CISRC-TR-72-14)

MATHIS, B. A. Techniques for the evaluation and improvement of computer-produced abstracts. December, 1972. 262p. (OSU-CISRC-TR-72-15) (PB-214 675)

1973

LAY, W. M. The double KWIC coordinate indexing technique: theory, design, and implementation. February, 1973. 250p. (OSU-CISRC-TR-73-1) (PB-229 499/AS)

YOUNG, C. E. Development of language andlysis procedures with application to automatic indexing. April, 1973. 298p. (OSU-CISRC-TR-73-2) (PB-227 088/AS)

STRONG, S. M. An algorithm for generating structual surrogates of English text. April, 1973. 140p. (OSU-CISRC-TR-73-3) (PB-227 071/AS)

HSIAO, D. K. Logical access control mechanisms in computer systems. July, 1973. 24p. (OSU-CISRC-TR-73-4)

WANG, P. T. R. Bandwidth minimization, reducibility decomposition, and triangularization of sparse matrices. September, 1973. 162p. (OSU-CISRC-TR-73-5)

BIERMANN, A. W.; BAUM, R.; KRISHNASWAMY, R.; PETRY, R. E. Automatic program synthesis reports. October, 1973. 62p. (OSU-CISRC-TR-73-6)

WYRICK, T.; HSIAO, D. K. System configuration study on the interconnection of the IBM 370/165, DEC System-10 and MICRO 1600/21 computer systems. October, 1973. 26p. (OSU-CISRC-TR-73-7)

1974

HSIAO, D. K.; KERR, D. S.; MC CAULEY III, E. J. A model for data secure systems (Part I). February, 1974. 39p. (OSU-CISRC-TR-73-8)

HSIAO, D. K.; KERR, D. S.; NEE, C. J. Context protection and consistent control in data base systems (Part I). February, 1974. 28p. (OSU-CISRC-TR-73-9)

BAUM, R. I.; HSIAO, D. K. A data secure computer architecture. July, 1974. 22p. (OSU-CISRC-TR-73-10)

RUSH, J. E.; PEPINSKY, H. B.; LANDRY, B. C.; MEARA, N. M.; STRONG, S. M.; VALLEY, J. A.; YOUNG, C. E. A computer assisted language analysis system. February, 1974. 71p. (OSU-CISRC-TR-74-1)

REAMES, C. C.; LIU, M. T. Variable-length message transmission for distributed loop computer networks (Part I). June, 1974. 9p. (OSU-CISRC-TR-74-2)

KIRSCH, B. M. An improved error diagnostics system for IBM System/360-370 assembler program dumps. June, 1974. 60p. (OSU-CISRC-TR-74-3)

COHEN, D.; LIU, M. T. Emulation of computer networks by microprogrammable microcomputers. June, 1974. 17p. (OSU-CISRC-TR-74-4)

INVESTIGATOR INDEX

Abilock, J. G., 19

Aitken, J., 14

Atwell, R. J., 46

Baum, R. 38

Beug, J. L., 17

Biermann, A. W., 27

Buttelmann, H. W., 21, 22, 23

Chandler, J. S., 12

Chandrasekaran, B., 25, 26, 27

Cheng, T. T., 49

Cohen, D., 47

Dasarathy, B., 34

DeLutis, T., 12, 13, 14

Dickey, F. J., 23

Doherty, M. E., 31

Ernst, R., 17

Farah, B., 48

Foulk, C. R., 43, 45

Fournier, S., 47

Friedman, M. S., 39

Gillenson, M. L., 27

Gudes, E., 37

Hasbrouck, C., 45

Hepler, S. P., 17

Horger, B., 44

Hsiao, D. K., 36, 38, 39, 40, 42, 44, 48

Hsu, J. H., 11

Jain, A. K., 25, 26

James, S. E., 14, 16

Jones, G. P., 16

Juelich, O. C., 43

Kaffen, N., 39

Kalmey, D. L., 32

Kammerer, B. J., 37, 44

Kar, G., 13, 16

Kerr, D. S., 31, 32, 41

Kirsch, B. M., 23, 41

Knablein, R., 37

Krishnaswamy, R., 23, 27

Koch, H. S., 37, 40, 42

Lam, C. C., 26

Lay, W. M., 11

Liu, M. T., 47, 48, 49

Longe, O., 45

Manola, r., 40, 42

Marik, D. A., 26

Mathis, R. F., 39, 41, 45, 46

McCauley, E. J., 39, 40

Moshell, M., 21

Muftic, S., 34

Nee, C. J., 36

Petrarca, A. E., 11, 12, 15

Petry, F. E., 27

Pumarada, A., 50

Pyster, A., 23, 24

Reames, C. C., 49

Rodeheffer, T., 36, 48

Rothstein, J., 21, 28, 30, 31, 33, 34, 35

Santhanam, V., 33

Sieberth, H., 37

Srihari, S. N., 25

Stahl, F., 37, 41, 42, 50

Stalcup, W. S., 11, 12, 15

Su, D. H. Y., 43

Wang, J. T., 48

Wang, P. T. R. 31

Westbrook, D. E., 16

White, L. J., 13, 16, 25, 26, 28, 31, 32, 33, 34

Wong, D., 34, 36

Young, W., 44

Yovits, M. C. 19

SUBJECT INDEX

Academic Programs, 4

Access and Access Control in Data Base Management Systems, A Model for, 42

Activities, Related, of CISRC Staff, 62

Aircraft Detection and Classification, 28

Algorithmic Control of Index Term Specificity, 11

Architecture, Data Secure System, 38

Artificial Intelligence, 25, 30

Attribute-Based File Organizations, 41

Automata on Unrestricted Syntactic Structures, Generalized Finite, 22

Automatic Document Classification, 13, 14, 16

Automatic Indexing, 11

Automatic Information Systems Analysis through Evaluation of Simulated Performance, 12

Automatic Program Synthesis, 27

Automatic Vocabulary Control in Printed Indexes, 12

Automatic Vocabulary Control Techniques, Evaluation of, 15

Binary Strings and Topology, 31

Cellular Automata and Parallel Computation, 34

Coding, Minimum-Redundancy Variable Length, 33

Complex File Structures, 44

Computer and Information Science Department, 4

Computer Communications, Mixed Data/Voice Transmission System for, 48

Computer Network, Local, 48

Computer Networks, 47, 48, 49

Computer Program Debugging, 46

Computer Program Flowgraphs, a Smoothness Index for, 45

Computer Science Education in Latin America, Investigation of, 50

Context Protection, 36

Control in Data Secure Systems, 36

Controlling Computer Program-Postmortem Diagnostic Output, 39

Course Offerings, 7, 51

Cross-Assembler for the Microdata 1600/21 Computer System, 36

Cryptography and Security Structures, 37, 41

Cubic Spline Approximations, Optimal Placement of Knots for, 33

Data Base, Graphical Model of Functional Dependences in a Relational, 41

Data Base Management, Performance Evaluation, 14

Data Base Management System, 37

Data Base Management Systems, Access and Access Control in, 42

Data Base Protection Mechanism, 39

Data Base Reorganization, the Development of Guidelines, 40

Data Definition, Mapping Functions for Independent, 42

Data Secure System Architecture, 38

Data Secure Systems, 36, 37, 39

Debugging System, Postmortem, 46

Debugging, Teaching, 45

Decision Theoretic Approach to Automatic Document Classification, 13

Decision Theoretic Models for Invariant Pattern Recognition, 25

Decision Theory, Finite Memory, 26

Decision Theory for Pattern Recognition, 25

Distance Functions for Independent Measurements and Finite Sample Size, 26

Doctoral Degrees, 9

Document Classification, Automatic, 13,14, 16

Double-KWIC Coordinate Indexing, 11

Equations, Comparison of Routines for Solving Nonlinear Systems of, 32

Error Diagnostics for IBM System /370 PS/MVT Assembler Program Dumps, 41

Evaluation of Automatic Vocabulary Control Techniques, 15

Facial Image on CRT, 27

Facilities of CISRC, 2

Faculty, 8, 55

File Partitioning in Attribute-Based File Organizations, 41

Finite Automata on Unrestricted Syntactic Structures, Generalized, 22

Finite Memory Decision Theory, 26

Finite Sample Size, 26

Flowgraphs, a Smoothness Index for Computer Program, 45

Formal Languages, 21, 28

Functional Dependences in a Relational Data Base, Graphical Model of, 41

Generalized Finite Automata on Unrestricted Syntactic Structures, 22

General Theory of Information Flow and Analysis, 19

Graduate Programs, 6

Grammars, Human Extrapolation of Strings Generated by

Ordered Cyclic Finite State, 17

Grammars, Syntactic Structures of Unrestricted, 21

Grammatical Inference Problem, 26

Graphical Model of Functional Dependencies in a Relational Data Base, 41

Groupoid String Algorithms, Turing Universality and Parallelism of, 35

Growth of Department of Computer and Information Science, 5

Human Extrapolation of Strings Generated by Ordered Cyclic Finite State Grammars, 17

Human Facial Image on CRT, 27

Human Information Processing, 17

Human Performance, Use of Probabilistic Automata as Models of, 17

IBM System /370 OS/MVT Assembler Program Dumps, 41

Independent Measurements and Finite Sample Size, Distance Functions for, 26

Indexing, Automatic, 11

Index Term Specificity in Double-KWIC Coordinate Indexing, 11

Information Analysis, 19

Information Flow and Analysis, A General Theory of, 17

Information Processes in Physical, Biological, and Social Systems, 30

Information Processing, Human, 17

Information Processing Systems Simulation, 13

Information Storage and Retrieval, 11

Information Systems Analysis Through Evaluation of Simulated Performance, 12

Keyterms in Automatic Document Classification, N-Gram, 14

Knots for Cubic Spline Approximations, Optimal Placement of, 34

Language Generation, Role of the Nonterminal in, 24

Language, Sieves and Context Sensitive, 33

Languages, Methods for Specifying Semantics of Context Free, 23

Languages, Syntax, Semantics and Translation of Phrase-Structure, 22

Linguistic Analysis, 21

Local Computer Network, 48

L^p-Norm Maxmin Problems, 33

Mapping Functions for Independent Data Definition, 42

Mathematical Techniques, 31

Matrices, Sparce, 31

Measurements and Finite Sample Size, Distance Functions for Independent, 26

Mechanized Information Center, 3

Memory, Finite, 26

MIC, 3

Microdata 1600/21 Computer System, 36

Networks, Computer, 47, 48, 49

N-Gram Keyterms in Automatic Document Classification, 14

Nonlinear Systems of Equations, Comparison of Routines for Solving, 32

Nonterminal in Language Generation, Role of the, 24

Number Trees, Semigroups, and Formal Languages, 28

Objectives of CISRC, 1

Ohio College Library Center (OCLC), 3

Optimization for Virtual Memory, Program, 26

Organization of CISRC, 1

Pagination of Programs for Virtual Memory Systems, 43

Parallel Computation, 34

Parallelism in Computer Programs, 43

Parallel Recognition of Formal Languages, 21

Pattern Recognition, Decision Theoretic Models for Invariant, 25

Pattern Recognition Decision Theory for, 25

Pattern Recognition Techniques and Computer Security, 42

Performance Evaluation of Data Base Management, 14

Phrase-Structure Languages, 22, 23

PL/X Systems Programming Language, 44

Privacy, Computer, 41

Probabilistic Automata as Models of Human Performance, Use of, 17

Properties of Syntax-Directed Translations, 23

Publications of CISRC Staff, 69

Quantitative Measure of Concept Scattering for Evaluation, 15

Radar Aircraft Detection and Classification, Low-Frequency Radar, 28

Record Placement in Attribute-Based File Organizations, 41

Reports, Technical, 73

Retrieval, Information Storage and, 11

Scope of Computer and Information Science Program, 2

Secure Systems, Data, 36, 37, 39

Secure Systems Model, A Data, 39

Security, Pattern Recognition Techniques and Computers, 42

Security Structures, Cryptography, and 23

Semantic-Directed Translation of Phrase Structure Languages, 22

Semantics of Context-Free Languages, Methods for Specifying, 23

Semigroups, and Formal Languages; Number Trees, 28

Seminar Series, Computer and Information Science, 59

Sieves and Context Sensitive Language, 33

Simulation, Information Processing Systems, 13

Smoothness Index for Computer Program Flowgraphs, 45

Sparse Matrices, 31

Storage and Retrieval, Information, 11

Sub-Optimal Set Covers with Confidence Measures of Optimality, 31

Syntactic Structures, Generalized Finite Automata on Unrestricted, 22

Syntactic Structures of Unrestricted Grammars, 21

Syntax-Directed and Semantic-Directed Translation of Phrase Structure Languages, 22

Syntax-Directed Translations, Properties of, 23

Syntax, Semantics and Translation of Phrase-Structure Languages, 22

Systems Programming, 36

Technical Reports, 9, 73

Turing Universality of Groupoid String Algorithms, 35

Topology, Binary Strings and, 31

Undergraduate Programs, 6

Virtual Memory Systems, 43

Vocabulary Control in Printed Indexes, Automatic, 12